GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 26, 2001, 08:20:53 ; Search time 21.15 Seconds
(without alignments)
960.238 Million cell updates/sec

Title: US-09-325-019-2
Perfect score: 1937
Sequence: 1 DFTPAPLEDTSSRPQFCKWP.....NPNDIFADLESYPDFSEIAN 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0601:*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1937	100.0	335	21	AAY59247	Human connective t
2	1937	100.0	345	20	AAY17640	Human putative mat
3	1937	100.0	367	20	AAY17641	Human WISP-1 prote
4	1937	100.0	367	22	AAB50975	Human PRO542 prote
5	1936	99.9	345	20	AAY17642	Human WISP-1 varia
6	1936	99.9	367	20	AAY17644	Human WISP-1 varia
7	1934	99.8	345	20	AAY17643	Human WISP-1 varia
8	1934	99.8	367	20	AAY17645	Human WISP-1 varia
9	1933	99.8	345	20	AAY17652	Human WISP-1 varia
10	1933	99.8	367	20	AAY17653	Human WISP-1 varia
11	1675	86.5	345	20	AAY17646	Mouse putative mat
12	1675	86.5	367	20	AAY17647	Mouse WISP-1 prote
13	869.5	44.9	349	16	AAR79964	Connective tissue
14	869.5	44.9	349	18	AAW09089	Human connective t
15	869.5	44.9	349	18	AAW11302	Connective tissue
16	869.5	44.9	349	19	AAW62084	Human connective t
17	869.5	44.9	349	20	AAY18361	Human connective t
18	869.5	44.9	349	20	AAW81425	Connective tissue
19	869.5	44.9	349	21	AAY92939	Human connective t
20	869.5	44.9	349	21	AAY92940	Human connective t
21	869.5	44.9	349	21	AAY44755	Human connective t
22	869.5	44.9	349	22	AAB60664	Human connective t
23	869.5	44.9	349	22	AAB48831	Human connective t
24	868.5	44.8	347	18	AAW12694	Connective tissue
25	846	43.7	339	20	AAY17654	Human putative mat
26	846	43.7	372	20	AAY17655	Human WISP-3 prote
27	844	43.6	339	20	AAY17656	Human putative mat
28	844	43.6	354	20	AAY34190	Human connective t
29	844	43.6	354	20	AAY17657	Human WISP-3 prote
30	844	43.6	354	21	AAY81438	Human growth facto
31	835.5	43.1	347	20	AAY24379	Rat connective tis
32	835.5	43.1	348	13	AAR25566	Beta-IG-M2. Mus m
33	831.5	42.9	348	18	AAW35731	Murine Fispl2. Mu
34	831.5	42.9	348	21	AAY44756	Mouse connective t
35	818	42.2	347	21	AAY93340	Amino acid sequenc
36	793.5	41.0	349	21	AAY44754	Bovine connective
37	772.5	39.9	379	13	AAR25565	Beta-IG-M1. Mus m
38	764.5	39.5	381	18	AAW35957	Human monocyte mat
39	764.5	39.5	455	21	AAB43987	Human cancer assoc
40	760.5	39.3	381	18	AAW35730	Human cysteine ric
41	743.5	38.4	351	14	AAR31599	Chicken nov protei
42	705.5	36.4	375	17	AAR90919	Connective tissue
43	705.5	36.4	375	20	AAY31620	Human CTGF-2. Hom
44	498.5	25.7	227	20	AAY27440	Rat HICP mature po
45	498.5	25.7	250	20	AAY27434	Rat HICP polypepti

ALIGNMENTS

RESULT 1
AAY59247
ID AAY59247 standard; Protein; 335 AA.
XX
AC AAY59247;
XX
DT 11-APR-2000 (first entry)
XX
DE Human connective tissue growth factor-4 (CTGF-4).
XX
KW Connective tissue growth factor-4; CTGF-4; human; immune system disorder;
KW hematopoietic disorder; autoimmune disorder; diabetes mellitus; asthma;
KW respiratory disorder; inflammation; hyperproliferative disorder;
KW infection; central nervous system disease; Alzheimer's disease; AIDS;
KW food additive.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 15..84
FT /note= "IGF binding domain"
FT Domain 28..36
FT /note= "conserved domain CD-I"
FT Domain 39..55
FT /note= "conserved domain CD-II"
FT Modified-site 54
FT /note= "potential N-glycosylation site"
FT Domain 61..70
FT /note= "conserved domain CD-III"
FT Domain 89..154
FT /note= "Von-Willebrand factor type C repeat fragment"
FT Domain 101..121
FT /note= "conserved domain CD-IV"
FT Modified-site 111

FT /note= "potential N-glycosylation site"
FT Domain 144-154
FT /note= "conserved domain CD-V"
FT Domain 184..228
FT /note= "sulfated glycoconjugate binding motif"
FT Domain 194..213
FT /note= "conserved domain CD-VI"
FT Domain 216..227
FT /note= "conserved domain CD-VII"
FT Domain 236..241
FT /note= "conserved domain CD-VIII"
FT Domain 241..316
FT /note= "C-terminal dimerisation and receptor-binding
FT domain"
FT Modified-site 252
FT /note= "potential N-glycosylation site"
FT Domain 253..260
FT /note= "conserved domain CD-IX"
FT Domain 264..280
FT /note= "conserved domain CD-X"
FT Domain 290..295
FT /note= "conserved domain CD-XI"
FT Modified-site 311
FT /note= "potential N-glycosylation site"
XX
PN WO9962927-A1.
XX
PD 09-DEC-1999.
XX
PF 03-JUN-1999; 99WO-US12150.
XX
PR 05-JUN-1998; 98US-0088320.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Young PE;
XX
DR WPI; 2000-147042/13.
DR N-PSDB; AAZ58613.
XX
PT New isolated connective tissue growth factor-4, used for treating e.g.
PT cancers -
XX
PS Claim 11; Fig 1A-E; 196pp; English.
XX
CC The invention provides an isolated human connective tissue growth factor
CC -4 (CTGF-4) polypeptide. The CTGF-4 cDNA is deposited under ATCC No.
CC 209816. The CTGF-4 protein can be expressed by standard recombinant
CC methodology. The polypeptides can be used for preventing, treating or
CC ameliorating a medical condition . They may be useful in treating
CC deficiencies or disorders of the immune system, by activating or
CC inhibiting the proliferation, differentiation, or mobilization
CC (chemotaxis) of immune cells, treating or detecting deficiencies or
CC disorders of hematopoietic cells (e.g. blood protein disorders, ataxia
CC telangiectasia, HIV infection, Di-george syndrome, anemia or
CC hemoglobinuria), to modulate hemostatic (the stopping of bleeding) or
CC thrombolytic activity (clot formation) (e.g. blood coagulation disorders,
CC blood platelet disorders, or wounds resulting from trauma, or surgery),
CC in treating or detecting autoimmune disorders (e.g. Addison's disease,
CC rheumatoid arthritis, allergic encephalomyelitis, Goodpastures syndrome,
CC multiple sclerosis, purpura, Reiter's disease, Guillain-Barre syndrome,
CC systemic lupus erythematosus, insulin dependent diabetes mellitus or
CC autoimmune inflammatory eye disease), treating asthma (particularly
CC allergic asthma) or other respiratory problems (e.g. anaphylaxis,
CC hypersensitivity to an antigenic molecule or blood group
CC incompatibility), to treat and/or prevent organ rejection or graft-versus
CC -host disease (GVHD), to modulate inflammation (septic shock, sepsis,
CC arthritis, nephritis, cytokine or chemokine induced lung injury,
CC inflammatory bowel disease, Crohn's disease, or resulting from over
CC production of cytokines), to treat hyperproliferative disorders,
CC including neoplasms in the abdomen, bone, breast, digestive system,
CC liver, pancreas, peritoneum, endocrine glands, eye, head and neck,
CC nervous (central and peripheral), lymphatic system, pelvic, skin, soft
CC tissue, spleen, thoracic and urogenital, hypergammaglobulinemia,
CC lymphoproliferative disorders, Waldenstron's macroglobulinemia,
CC sarcoidosis), to treat or detect infectious agents, e.g. viruses (e.g.
CC arthritis, bronchiollitis, encephalitis, eye infections, chronic fatigue
CC syndrome, hepatitis, meningitis, AIDS, pneumonia, chickenpox, measles,
CC mumps, parainfluenza, rabies, the common cold, polio, leukemia, rubella,
CC sexually transmitted diseases, or skin diseases) bacterial or fungal
CC agents (e.g. bacteremia, endocarditis, eye infections, gingivitis,
CC opportunistic infections, respiratory tract infections, Lyme disease,
CC cat-scratch disease, paratyphoid fever, food poisoning, pneumonia,
CC gonorrhea and sexually transmitted diseases, meningitis, tuberculosis,
CC lupus, gangrene, tetanus, rheumatic fever, urinary tract infections,
CC wound infections), parasitic agents (e.g. scabies, dysentery, liver
CC disease, malaria, toxoplasmosis), to differentiate, proliferate and
CC attract cells, leading to the regeneration of tissues (e.g. repair,
CC replace or protect tissue in wounds, burns, incisions or ulcers,
CC osteoporosis, osteocarthritis, periodontal disease, liver failure,
CC surgery, cosmetic plastic surgery, reperfusion injury) to proliferate and
CC differentiate nerve cells (e.g. spinal cord disorders, head trauma,
CC cerebrovascular disease and stroke), localized neuropathies and central
CC nervous system diseases (e.g. Alzheimer's disease, Parkinson's disease,
CC Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager
CC syndrome). They may also increase or decrease the differentiation or
CC proliferation of embryonic stem cells and hematopoietic lineage, may be
CC used to modulate mammalian characteristics such as body height, weight,
CC hair color, eye color, skin, percentage of adipose tissue, pigmentation,
CC size, and shape, to modulate mammalian metabolism affecting catabolism,
CC anabolism, processing, utilization and storage of energy, to change a
CC mammal's mental state or physical state by influencing biorhythms,
CC caricadic rhythms, circadian rhythms, depression (including depressive
CC disorders), tendency for violence, tolerance for pain, reproductive
CC capabilities, hormonal or endocrine levels, appetite, libido, memory,
CC stress, or other cognitive qualities, as a food additive or preservative,
CC such as to increase or decrease storage capabilities, fat content, lipid,
CC protein, carbohydrate, vitamins, minerals, cofactors or other nutritional
CC components. Mutations in the PNs or the presence or amount of expression
CC or activity of the polypeptides can be used for diagnosing a pathological
CC condition or a susceptibility to a pathological condition. The CFTR-4
CC polypeptides can also be used for identifying binding partners . The
CC products can also be used for producing transgenic animals. The present
CC sequence represents the CTGF-4 polypeptide.
XX
SQ Sequence 335 AA;

Query Match 100.0%; Score 1937; DB 21; Length 335;
Best Local Similarity 100.0%; Pred. No. 5.5e-138;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 dftpapledtssrpqfckwpcecppspprcplgvslitdgcecckmcaqqlgdnctеaai 60

Qy     61 CDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 cdphrglycdysgdrpryaigvcaqvvgvgcvldgvrynngqsfqpnckynctcidgavg 120

Qy    121 CTPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAW 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ctplclrvrpprlwcphprrvsipghcceqwvceddakrprktaprdtgafdavgeveaw 180

Qy    181 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 hrnciaytspwspcstscglgvstrisnvnaqcwpeqesrlcnlrpcdvdihtlikagkk 240

Qy    241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 clavyqpeasmnftlagcistrsyqpkycgvcmdnrccipyksktidvsfqcpdglgfsr 300

Qy    301 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
          |||||||||||||||||||||||||||||||||||
Db    301 qvlwinacfcnlscrnpndifadlesypdfseian 335
```

```
RESULT   2
AAY17640
ID   AAY17640 standard; Protein; 345 AA.
XX
AC   AAY17640;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Human putative mature WISP-1 protein SEQ ID NO:3.
XX
KW   WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW   connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW   leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW   tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW   kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW   connective tissue disorder; catabolic state; inflammation;
KW   testicular-related disorder; angiogenesis; immunological disorder.
XX
OS   Homo sapiens.
XX
PN   WO9921998-A1.
XX
PD   06-MAY-1999.
XX
PF   29-OCT-1998;   98WO-US22991.
XX
PR   14-APR-1998;   98US-0081695.
PR   29-OCT-1997;   97US-0063704.
PR   03-FEB-1998;   98US-0073612.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein DA,  Cohen RL,  Goddard A,  Gurney AL,  Hillan K;
PI   Lawrence DA,  Levine AJ,  Pennica D,  Roy MA,  Wood WI;
XX
DR   WPI; 1999-337420/28.
XX
PT   New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS   Claim 1; Page 162-163; 284pp; English.
XX
CC   The present invention describes Wnt-1 induced secreted polypeptides,
CC   WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC   and WISP-3 have homology to connective tissue growth factor (CTGF).
CC   Products from the present invention can be used to treat WISP-related
CC   disorders such as breast, ovarian, and colon cancer or melanoma. The
CC   products can be used to treat arteriosclerosis. The products can also be
CC   used to treat other diseases e.g. benign and malignant tumours,
CC   leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC   blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC   disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC   disorders, bone-related disorders such as osteoporosis, trauma such as
CC   burns, incisions, and other wounds, connective tissue disorders,
CC   catabolic states, testicular-related disorders, and inflammatory,
CC   angiogenic and immunologic disorders including arteriosclerosis. The
CC   products can also be used for detection and diagnosis especially of
CC   individuals with neoplastic cell growth or proliferation. The products
CC   can be used in the production of transgenic or knock-out animals.
CC   Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC   cells.
XX
SQ   Sequence   345 AA;


  Query Match           100.0%;  Score 1937;  DB 20;  Length 345;
  Best Local Similarity  100.0%;  Pred. No. 5.7e-138;
  Matches  335;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11 dftpapledtssrpqfckwpcecppspprcplgvslitdgcecckmcaqqlgdncteaai 70

Qy      61 CDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      71 cdphrglycdysgdrpryaigvcaqvvgvgcvldgvrynngqsfqpnckynctcidgavg 130

Qy     121 CTPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAW 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     131 ctplclrvrpprlwcphprrvsipghcceqwvceddakrprktaprdtgafdavgeveaw 190

Qy     181 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     191 hrnciaytspwspcstscglgvstrisnvnaqcwpeqesrlcnlrpcdvdihtlikagkk 250

Qy     241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251 clavyqpeasmnftlagcistrsyqpkycgvcmdnrccipyksktidvsfqcpdglgfsr 310

Qy     301 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
           ||||||||||||||||||||||||||||||||||
Db     311 qvlwinacfcnlscrnpndifadlesypdfseian 345


RESULT   3
AAY17641
ID   AAY17641 standard; Protein; 367 AA.
XX
AC   AAY17641;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Human WISP-1 protein SEQ ID NO:4.
XX
KW   WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW   connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW   leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW   tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW   kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW   connective tissue disorder; catabolic state; inflammation;
KW   testicular-related disorder; angiogenesis; immunological disorder.
XX
OS   Homo sapiens.
XX
PN   WO9921998-A1.
XX
PD   06-MAY-1999.
XX
PF   29-OCT-1998;   98WO-US22991.
XX
PR   14-APR-1998;   98US-0081695.
PR   29-OCT-1997;   97US-0063704.
PR   03-FEB-1998;   98US-0073612.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein DA,  Cohen RL,  Goddard A,  Gurney AL,  Hillan K;
PI   Lawrence DA,  Levine AJ,  Pennica D,  Roy MA,  Wood WI;
XX
DR   WPI; 1999-337420/28.
DR   N-PSDB; AAX76482.
XX
PT   New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS   Claim 4; Page 163-164; 284pp; English.
XX
CC   The present invention describes Wnt-1 induced secreted polypeptides,
CC   WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC   and WISP-3 have homology to connective tissue growth factor (CTGF).
CC   Products from the present invention can be used to treat WISP-related
CC   disorders such as breast, ovarian, and colon cancer or melanoma. The
CC   products can be used to treat arteriosclerosis. The products can also be
CC   used to treat other diseases e.g. benign and malignant tumours,
CC   leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal, and
```

CC blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC disorders, bone-related disorders such as osteoporosis, trauma such as
CC burns, incisions, and other wounds, connective tissue disorders,
CC catabolic states, testicular-related disorders, and inflammatory,
CC angiogenic and immunologic disorders including arteriosclerosis. The
CC products can also be used for detection and diagnosis especially of
CC individuals with neoplastic cell growth or proliferation. The products
CC can be used in the production of transgenic or knock-out animals.
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC cells.
XX
SQ Sequence 367 AA;

Query Match 100.0%; Score 1937; DB 20; Length 367;
Best Local Similarity 100.0%; Pred. No. 6.1e-138;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      33 dftpapledtssrpqfckwpcecppspprcplgvslitdgcecckmcaqqlgdncteaai 92

Qy      61 CDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93 cdphrglycdysgdrpryaigvcaqvvgvgcvldgvrynngqsfqpnckynctcidgavg 152

Qy     121 CTPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAW 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     153 ctplclrvrpprlwcphprrvsipghcceqwvceddakrprktaprdtgafdavgeveaw 212

Qy     181 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     213 hrnciaytspwspcstscglgvstrisnvnaqcwpeqesrlcnlrpcdvdihtlikagkk 272

Qy     241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     273 clavyqpeasmnftlagcistrsyqpkycgvcmdnrccipyksktidvsfqcpdglgfsr 332

Qy     301 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
          |||||||||||||||||||||||||||||||||||
Db     333 qvlwinacfcnlscrnpndifadlesypdfseian 367
```

RESULT 4
AAB50975
ID AAB50975 standard; Protein; 367 AA.
XX
AC AAB50975;
XX
DT 21-MAR-2001 (first entry)
XX
DE Human PRO542 protein.
XX
KW Human; PRO; cytostatic; nootropic; neuroprotective; respiratory general;
KW antiinflammatory; antiangiogenic; immunosuppressive; immunostimulant;
KW PRO agonist; cancer; inflammatory disorder; immunological disorder.
XX
OS Homo sapiens.
XX
PN WO200073348-A2.
XX
PD 07-DEC-2000.
XX
PF 30-MAY-2000; 2000WO-US14941.
XX
PR 02-JUN-1999; 99WO-US12252.
PR 22-JUN-1999; 99US-0140650.
PR 23-JUN-1999; 99US-0141037.
PR 20-JUL-1999; 99US-0144758.
PR 01-SEP-1999; 99WO-US20111.
PR 08-SEP-1999; 99WO-US20594.
PR 29-OCT-1999; 99US-0162506.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30999.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 02-MAR-2000; 2000WO-US05841.
PR 03-MAR-2000; 2000US-0187202.
PR 10-MAR-2000; 2000WO-US06319.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W, Kabakoff RC;
PI Shelton DL, Smith V, Watanabe CK, Wood WI;
XX
DR WPI; 2001-016509/02.
DR N-PSDB; AAC91577.
XX
PT Twenty eight nucleic acids encoding PRO polypeptides which are useful
PT for treating various tumors, e.g. breast cancer, and other
PT inflammatory, angiogenic and immunological disorders -
XX
PS Claim 31; Fig 50; 188pp; English.
XX
CC The present sequence is one of twenty eight novel PRO polypeptides. The
CC PRO polypeptides and their agonists, including antibodies, peptides, and
CC small molecule agonists, may be used to treat various tumours, e.g.,
CC cancers such as breast cancer, ovarian cancer, renal cancer, colorectal
CC cancer, uterine cancer, prostate cancer, lung cancer, bladder cancer,
CC central nervous system cancer, melanoma or leukaemia. They are also
CC useful for treating other disorders such as neuronal, glial, astrocytal,
CC hypothalamic and other glandular, macrophagal, epithelial, stromal and
CC blastocoelic disorders, and inflammatory, angiogenic and immunological
CC disorders.
XX
SQ Sequence 367 AA;

Query Match 100.0%; Score 1937; DB 22; Length 367;
Best Local Similarity 100.0%; Pred. No. 6.1e-138;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      33 dftpapledtssrpqfckwpcecppspprcplgvslitdgcecckmcaqqlgdncteaai 92

Qy      61 CDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93 cdphrglycdysgdrpryaigvcaqvvgvgcvldgvrynngqsfqpnckynctcidgavg 152

Qy     121 CTPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAW 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     153 ctplclrvrpprlwcphprrvsipghcceqwvceddakrprktaprdtgafdavgeveaw 212

Qy     181 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     213 hrnciaytspwspcstscglgvstrisnvnaqcwpeqesrlcnlrpcdvdihtlikagkk 272

Qy     241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     273 clavyqpeasmnftlagcistrsyqpkycgvcmdnrccipyksktidvsfqcpdglgfsr 332

Qy     301 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
          |||||||||||||||||||||||||||||||||||
Db     333 qvlwinacfcnlscrnpndifadlesypdfseian 367
```

```
RESULT   5
AAY17642
ID   AAY17642 standard; Protein; 345 AA.
XX
AC   AAY17642;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Human WISP-1 variant protein SEQ ID NO:5.
XX
KW   WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW   connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW   leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW   tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW   kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW   connective tissue disorder; catabolic state; inflammation;
KW   testicular-related disorder; angiogenesis; immunological disorder.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9921998-A1.
XX
PD   06-MAY-1999.
XX
PF   29-OCT-1998;   98WO-US22991.
XX
PR   14-APR-1998;   98US-0081695.
PR   29-OCT-1997;   97US-0063704.
PR   03-FEB-1998;   98US-0073612.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein DA,  Cohen RL,  Goddard A,  Gurney AL,  Hillan K;
PI   Lawrence DA,  Levine AJ,  Pennica D,  Roy MA,  Wood WI;
XX
DR   WPI; 1999-337420/28.
XX
PT   New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS   Claim 5; Page 164-165; 284pp; English.
XX
CC   The present invention describes Wnt-1 induced secreted polypeptides,
CC   WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC   and WISP-3 have homology to connective tissue growth factor (CTGF).
CC   Products from the present invention can be used to treat WISP-related
CC   disorders such as breast, ovarian, and colon cancer or melanoma. The
CC   products can be used to treat arteriosclerosis. The products can also be
CC   used to treat other diseases e.g. benign and malignant tumours,
CC   leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC   blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC   disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC   disorders, bone-related disorders such as osteoporosis, trauma such as
CC   burns, incisions, and other wounds, connective tissue disorders,
CC   catabolic states, testicular-related disorders, and inflammatory,
CC   angiogenic and immunologic disorders including arteriosclerosis. The
CC   products can also be used for detection and diagnosis especially of
CC   individuals with neoplastic cell growth or proliferation. The products
CC   can be used in the production of transgenic or knock-out animals.
CC   Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC   cells.
XX
SQ   Sequence   345 AA;


  Query Match           99.9%;  Score 1936;  DB 20;  Length 345;
  Best Local Similarity  99.7%;  Pred. No. 6.8e-138;
  Matches  334;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11 dftpapledtssrpqfckwpcecppspprcplgvslitdgcecckmcaqqlgdncteaai 70

Qy      61 CDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      71 cdphrglycdysgdrpryaigvcaqvvgvgcvldgvrynngqsfqpnckynctcidgavg 130

Qy     121 CTPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAW 180
           ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db     131 ctplclrvrpprlwcphprrvsipghcceqwiceddakrprktaprdtgafdavgeveaw 190

Qy     181 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     191 hrnciaytspwspcstscglgvstrisnvnaqcwpeqesrlcnlrpcdvdihtlikagkk 250

Qy     241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251 clavyqpeasmnftlagcistrsyqpkycgvcmdnrccipyksktidvsfqcpdglgfsr 310

Qy     301 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
           |||||||||||||||||||||||||||||||||||
Db     311 qvlwinacfcnlscrnpndifadlesypdfseian 345


RESULT   6
AAY17644
ID   AAY17644 standard; Protein; 367 AA.
XX
AC   AAY17644;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Human WISP-1 variant protein SEQ ID NO:7.
XX
KW   WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW   connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW   leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW   tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW   kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW   connective tissue disorder; catabolic state; inflammation;
KW   testicular-related disorder; angiogenesis; immunological disorder.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9921998-A1.
XX
PD   06-MAY-1999.
XX
PF   29-OCT-1998;   98WO-US22991.
XX
PR   14-APR-1998;   98US-0081695.
PR   29-OCT-1997;   97US-0063704.
PR   03-FEB-1998;   98US-0073612.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein DA,  Cohen RL,  Goddard A,  Gurney AL,  Hillan K;
PI   Lawrence DA,  Levine AJ,  Pennica D,  Roy MA,  Wood WI;
XX
DR   WPI; 1999-337420/28.
XX
PT   New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS   Claim 6; Page 167-168; 284pp; English.
XX
CC   The present invention describes Wnt-1 induced secreted polypeptides,
CC   WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC   and WISP-3 have homology to connective tissue growth factor (CTGF).
CC   Products from the present invention can be used to treat WISP-related
CC   disorders such as breast, ovarian, and colon cancer or melanoma. The
CC   products can be used to treat arteriosclerosis. The products can also be
CC   used to treat other diseases e.g. benign and malignant tumours,
CC   leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal, and
```

```
CC   blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC   disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC   disorders, bone-related disorders such as osteoporosis, trauma such as
CC   burns, incisions, and other wounds, connective tissue disorders,
CC   catabolic states, testicular-related disorders, and inflammatory,
CC   angiogenic and immunologic disorders including arteriosclerosis. The
CC   products can also be used for detection and diagnosis especially of
CC   individuals with neoplastic cell growth or proliferation. The products
CC   can be used in the production of transgenic or knock-out animals.
CC   Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC   cells.
XX
SQ   Sequence   367 AA;


  Query Match            99.9%;  Score 1936;  DB 20;  Length 367;
  Best Local Similarity  99.7%;  Pred. No. 7.2e-138;
  Matches  334;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     33 dftpapledtssrpqfckwpcecppspprcplgvslitdgcecckmcaqqlgdncteaai 92

Qy     61 CDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     93 cdphrglycdysgdrpryaigvcaqvvgvgcvldgvrynngqsfqpnckynctcidgavg 152

Qy    121 CTPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAW 180
          ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db    153 ctplclrvrpprlwcphprrvsipghcceqwiceddakrprktaprdtgafdavgeveaw 212

Qy    181 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    213 hrnciaytspwspcstscglgvstrisnvnaqcwpeqesrlcnlrpcdvdihtlikagkk 272

Qy    241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    273 clavyqpeasmnftlagcistrsyqpkycgvcmdnrccipyksktidvsfqcpdglgfsr 332

Qy    301 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
          |||||||||||||||||||||||||||||||||||
Db    333 qvlwinacfcnlscrnpndifadlesypdfseian 367


RESULT   7
AAY17643
ID   AAY17643 standard; Protein; 345 AA.
XX
AC   AAY17643;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Human WISP-1 variant protein SEQ ID NO:6.
XX
KW   WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW   connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW   leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW   tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW   kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW   connective tissue disorder; catabolic state; inflammation;
KW   testicular-related disorder; angiogenesis; immunological disorder.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9921998-A1.
XX
PD   06-MAY-1999.
XX
PF   29-OCT-1998;   98WO-US22991.
XX
PR   14-APR-1998;   98US-0081695.
PR   29-OCT-1997;   97US-0063704.
PR   03-FEB-1998;   98US-0073612.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein DA,  Cohen RL,  Goddard A,  Gurney AL,  Hillan K;
PI   Lawrence DA,  Levine AJ,  Pennica D,  Roy MA,  Wood WI;
XX
DR   WPI; 1999-337420/28.
XX
PT   New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS   Claim 5; Page 166-167; 284pp; English.
XX
CC   The present invention describes Wnt-1 induced secreted polypeptides,
CC   WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC   and WISP-3 have homology to connective tissue growth factor (CTGF).
CC   Products from the present invention can be used to treat WISP-related
CC   disorders such as breast, ovarian, and colon cancer or melanoma. The
CC   products can be used to treat arteriosclerosis. The products can also be
CC   used to treat other diseases e.g. benign and malignant tumours,
CC   leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC   blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC   disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC   disorders, bone-related disorders such as osteoporosis, trauma such as
CC   burns, incisions, and other wounds, connective tissue disorders,
CC   catabolic states, testicular-related disorders, and inflammatory,
CC   angiogenic and immunologic disorders including arteriosclerosis. The
CC   products can also be used for detection and diagnosis especially of
CC   individuals with neoplastic cell growth or proliferation. The products
CC   can be used in the production of transgenic or knock-out animals.
CC   Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC   cells.
XX
SQ   Sequence   345 AA;


  Query Match            99.8%;  Score 1934;  DB 20;  Length 345;
  Best Local Similarity  99.7%;  Pred. No. 9.6e-138;
  Matches  334;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     11 dftpapledtssrpqfckwpcecppspprcplgvslitdgcecckmcaqqlgdncteaai 70

Qy     61 CDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     71 cdphrglycdysgdrpryaigvcaqvvgvgcvldgvrynngqsfqpnckynctcidgavg 130

Qy    121 CTPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAW 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db    131 ctplclrvrpprlwcphprrvsipghcceqwvceddakrprktaprdtgsfdavgeveaw 190

Qy    181 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    191 hrnciaytspwspcstscglgvstrisnvnaqcwpeqesrlcnlrpcdvdihtlikagkk 250

Qy    241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    251 clavyqpeasmnftlagcistrsyqpkycgvcmdnrccipyksktidvsfqcpdglgfsr 310

Qy    301 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
          |||||||||||||||||||||||||||||||||||
Db    311 qvlwinacfcnlscrnpndifadlesypdfseian 345


RESULT   8
AAY17645
ID   AAY17645 standard; Protein; 367 AA.
XX
AC   AAY17645;
XX
DT   06-AUG-1999  (first entry)
```

```
XX
DE   Human WISP-1 variant protein SEQ ID NO:8.
XX
KW   WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW   connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW   leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW   tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW   kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW   connective tissue disorder; catabolic state; inflammation;
KW   testicular-related disorder; angiogenesis; immunological disorder.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9921998-A1.
XX
PD   06-MAY-1999.
XX
PF   29-OCT-1998;   98WO-US22991.
XX
PR   14-APR-1998;   98US-0081695.
PR   29-OCT-1997;   97US-0063704.
PR   03-FEB-1998;   98US-0073612.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein DA,  Cohen RL,  Goddard A,  Gurney AL,  Hillan K;
PI   Lawrence DA,  Levine AJ,  Pennica D,  Roy MA,  Wood WI;
XX
DR   WPI; 1999-337420/28.
XX
PT   New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS   Claim 6; Page 168-169; 284pp; English.
XX
CC   The present invention describes Wnt-1 induced secreted polypeptides,
CC   WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC   and WISP-3 have homology to connective tissue growth factor (CTGF).
CC   Products from the present invention can be used to treat WISP-related
CC   disorders such as breast, ovarian, and colon cancer or melanoma. The
CC   products can be used to treat arteriosclerosis. The products can also be
CC   used to treat other diseases e.g. benign and malignant tumours,
CC   leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC   blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC   disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC   disorders, bone-related disorders such as osteoporosis, trauma such as
CC   burns, incisions, and other wounds, connective tissue disorders,
CC   catabolic states, testicular-related disorders, and inflammatory,
CC   angiogenic and immunologic disorders including arteriosclerosis. The
CC   products can also be used for detection and diagnosis especially of
CC   individuals with neoplastic cell growth or proliferation. The products
CC   can be used in the production of transgenic or knock-out animals.
CC   Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC   cells.
XX
SQ   Sequence   367 AA;


  Query Match             99.8%;  Score 1934;  DB 20;  Length 367;
  Best Local Similarity   99.7%;  Pred. No. 1e-137;
  Matches  334;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      33 dftpapledtssrpqfckwpcecppspprcplgvslitdgcecckmcaqqlgdncteaai 92

Qy      61 CDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93 cdphrglycdysgdrpryaigvcaqvvgvgcvldgvrynngqsfqpnckynctcidgavg 152

Qy     121 CTPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAW 180
           ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||

Db     153 ctplclrvrpprlwcphprrvsipghcceqwvceddakrprktaprdtgsfdavgeveaw 212

Qy     181 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     213 hrnciaytspwspcstscglgvstrisnvnaqcwpeqesrlcnlrpcdvdihtlikagkk 272

Qy     241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     273 clavyqpeasmnftlagcistrsyqpkycgvcmdnrccipyksktidvsfqcpdglgfsr 332

Qy     301 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
           |||||||||||||||||||||||||||||||||||
Db     333 qvlwinacfcnlscrnpndifadlesypdfseian 367


RESULT   9
AAY17652
ID   AAY17652 standard; Protein; 345 AA.
XX
AC   AAY17652;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Human WISP-1 variant protein SEQ ID NO:21.
XX
KW   WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW   connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW   leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW   tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW   kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW   connective tissue disorder; catabolic state; inflammation;
KW   testicular-related disorder; angiogenesis; immunological disorder.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9921998-A1.
XX
PD   06-MAY-1999.
XX
PF   29-OCT-1998;   98WO-US22991.
XX
PR   14-APR-1998;   98US-0081695.
PR   29-OCT-1997;   97US-0063704.
PR   03-FEB-1998;   98US-0073612.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein DA,  Cohen RL,  Goddard A,  Gurney AL,  Hillan K;
PI   Lawrence DA,  Levine AJ,  Pennica D,  Roy MA,  Wood WI;
XX
DR   WPI; 1999-337420/28.
XX
PT   New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS   Claim 7; Page 182-183; 284pp; English.
XX
CC   The present invention describes Wnt-1 induced secreted polypeptides,
CC   WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC   and WISP-3 have homology to connective tissue growth factor (CTGF).
CC   Products from the present invention can be used to treat WISP-related
CC   disorders such as breast, ovarian, and colon cancer or melanoma. The
CC   products can be used to treat arteriosclerosis. The products can also be
CC   used to treat other diseases e.g. benign and malignant tumours,
CC   leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC   blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC   disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC   disorders, bone-related disorders such as osteoporosis, trauma such as
CC   burns, incisions, and other wounds, connective tissue disorders,
CC   catabolic states, testicular-related disorders, and inflammatory,
CC   angiogenic and immunologic disorders including arteriosclerosis. The
CC   products can also be used for detection and diagnosis especially of
```

```
CC   individuals with neoplastic cell growth or proliferation. The products
CC   can be used in the production of transgenic or knock-out animals.
CC   Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC   cells.
XX
SQ   Sequence   345 AA;


  Query Match             99.8%;  Score 1933;  DB 20;  Length 345;
  Best Local Similarity   99.4%;  Pred. No. 1.1e-137;
  Matches  333;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11 dftpapledtssrpqfckwpcecppspprcplgvslitdgcecckmcaqqlgdncteaai 70

Qy      61 CDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      71 cdphrglycdysgdrpryaigvcaqvvgvgcvldgvrynngqsfqpnckynctcidgavg 130

Qy     121 CTPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAW 180
           ||||||||||||||||||||||||||||:||||||||||||||||:||||||||||
Db     131 ctplclrvrpprlwcphprrvsipghcceqwiceddakrprktaprdtgsfdavgeveaw 190

Qy     181 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     191 hrnciaytspwspcstscglgvstrisnvnaqcwpeqesrlcnlrpcdvdihtlikagkk 250

Qy     241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251 clavyqpeasmnftlagcistrsyqpkycgvcmdnrccipyksktidvsfqcpdglgfsr 310

Qy     301 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
           |||||||||||||||||||||||||||||||||||
Db     311 qvlwinacfcnlscrnpndifadlesypdfseian 345


RESULT  10
AAY17653
ID   AAY17653 standard; Protein; 367 AA.
XX
AC   AAY17653;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Human WISP-1 variant protein SEQ ID NO:22.
XX
KW   WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW   connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW   leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW   tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW   kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW   connective tissue disorder; catabolic state; inflammation;
KW   testicular-related disorder; angiogenesis; immunological disorder.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9921998-A1.
XX
PD   06-MAY-1999.
XX
PF   29-OCT-1998;   98WO-US22991.
XX
PR   14-APR-1998;   98US-0081695.
PR   29-OCT-1997;   97US-0063704.
PR   03-FEB-1998;   98US-0073612.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein DA,  Cohen RL,  Goddard A,  Gurney AL,  Hillan K;
PI   Lawrence DA,  Levine AJ,  Pennica D,  Roy MA,  Wood WI;
XX
DR   WPI; 1999-337420/28.
XX
PT   New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS   Claim 7; Page 183-184; 284pp; English.
XX
CC   The present invention describes Wnt-1 induced secreted polypeptides,
CC   WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC   and WISP-3 have homology to connective tissue growth factor (CTGF).
CC   Products from the present invention can be used to treat WISP-related
CC   disorders such as breast, ovarian, and colon cancer or melanoma. The
CC   products can be used to treat arteriosclerosis. The products can also be
CC   used to treat other diseases e.g. benign and malignant tumours,
CC   leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC   blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC   disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC   disorders, bone-related disorders such as osteoporosis, trauma such as
CC   burns, incisions, and other wounds, connective tissue disorders,
CC   catabolic states, testicular-related disorders, and inflammatory,
CC   angiogenic and immunologic disorders including arteriosclerosis. The
CC   products can also be used for detection and diagnosis especially of
CC   individuals with neoplastic cell growth or proliferation. The products
CC   can be used in the production of transgenic or knock-out animals.
CC   Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC   cells.
XX
SQ   Sequence   367 AA;


  Query Match             99.8%;  Score 1933;  DB 20;  Length 367;
  Best Local Similarity   99.4%;  Pred. No. 1.2e-137;
  Matches  333;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      33 dftpapledtssrpqfckwpcecppspprcplgvslitdgcecckmcaqqlgdncteaai 92

Qy      61 CDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93 cdphrglycdysgdrpryaigvcaqvvgvgcvldgvrynngqsfqpnckynctcidgavg 152

Qy     121 CTPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAW 180
           ||||||||||||||||||||||||||||:||||||||||||||||:||||||||||
Db     153 ctplclrvrpprlwcphprrvsipghcceqwiceddakrprktaprdtgsfdavgeveaw 212

Qy     181 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     213 hrnciaytspwspcstscglgvstrisnvnaqcwpeqesrlcnlrpcdvdihtlikagkk 272

Qy     241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     273 clavyqpeasmnftlagcistrsyqpkycgvcmdnrccipyksktidvsfqcpdglgfsr 332

Qy     301 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
           |||||||||||||||||||||||||||||||||||
Db     333 qvlwinacfcnlscrnpndifadlesypdfseian 367


RESULT  11
AAY17646
ID   AAY17646 standard; Protein; 345 AA.
XX
AC   AAY17646;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Mouse putative mature WISP-1 protein SEQ ID NO:11.
XX
KW   WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW   connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW   leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW   tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
```

```
KW   kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW   connective tissue disorder; catabolic state; inflammation;
KW   testicular-related disorder; angiogenesis; immunological disorder.
XX
OS   Mus sp.
XX
PN   WO9921998-A1.
XX
PD   06-MAY-1999.
XX
PF   29-OCT-1998;   98WO-US22991.
XX
PR   14-APR-1998;   98US-0081695.
PR   29-OCT-1997;   97US-0063704.
PR   03-FEB-1998;   98US-0073612.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein DA,  Cohen RL,  Goddard A,  Gurney AL,  Hillan K;
PI   Lawrence DA,  Levine AJ,  Pennica D,  Roy MA,  Wood WI;
XX
DR   WPI; 1999-337420/28.
XX
PT   New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS   Claim 9; Page 172-173; 284pp; English.
XX
CC   The present invention describes Wnt-1 induced secreted polypeptides,
CC   WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC   and WISP-3 have homology to connective tissue growth factor (CTGF).
CC   Products from the present invention can be used to treat WISP-related
CC   disorders such as breast, ovarian, and colon cancer or melanoma. The
CC   products can be used to treat arteriosclerosis. The products can also be
CC   used to treat other diseases e.g. benign and malignant tumours,
CC   leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC   blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC   disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC   disorders, bone-related disorders such as osteoporosis, trauma such as
CC   burns, incisions, and other wounds, connective tissue disorders,
CC   catabolic states, testicular-related disorders, and inflammatory,
CC   angiogenic and immunologic disorders including arteriosclerosis. The
CC   products can also be used for detection and diagnosis especially of
CC   individuals with neoplastic cell growth or proliferation. The products
CC   can be used in the production of transgenic or knock-out animals.
CC   Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC   cells.
XX
SQ   Sequence   345 AA;


  Query Match             86.5%;  Score 1675;  DB 20;  Length 345;
  Best Local Similarity   85.3%;  Pred. No. 2.6e-118;
  Matches  285;  Conservative   20;  Mismatches   29;  Indels    0;  Gaps    0;

Qy       2 FTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAIC 61
           |||||||:|::||:|||||||| ||||||||||||||||||:||||||||||||||
Db      12 ftpapleetttrpefckwpcecpqspprcplgvslitdgcecckicaqqlgdncteaaic 71

Qy      62 DPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGC 121
           |||||||||||||||||||||||||||||||||| ||:||||||:|||||||| |||
Db      72 dphrglycdysgdrpryaigvcaqvvgvgcvldgvrytngesfqpncrynctcidgtvgc 131

Qy     122 TPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWH 181
           |||||  |||||||  || | :|| |||||||:|||:|||:||  || || | | ||  :
Db     132 tplclsprppprlwcrqprhvrvpgqcceqwvcdddarrprqtalldtrafaasgaveqry 191

Qy     182 RNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKKC 241
            ||||||||||||:||||:|||||||||:|||||||||||||||||||   |||||||
Db     192 enciaytspwspcsttcglgistrisnvnarcwpeqesrlcnlrpcdvdiqlhikagkkc 251

Qy     242 LAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSRQ 301
           ||||||| : |||||||:|||:|:||||||| |||||||||||| | ||||:| |||||
Db     252 lavyqpeeatnftlagcvstrtyrpkycgvctdnrccipyksktisvdfqcpegpgfsrq 311

Qy     302 VLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
           |||||||||||||||||||||||||||||| ||||
Db     312 vlwinacfcnlscrnpndifadlesypdfeeian 345


RESULT  12
AAY17647
ID   AAY17647 standard; Protein; 367 AA.
XX
AC   AAY17647;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Mouse WISP-1 protein SEQ ID NO:12.
XX
KW   WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW   connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW   leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW   tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW   kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW   connective tissue disorder; catabolic state; inflammation;
KW   testicular-related disorder; angiogenesis; immunological disorder.
XX
OS   Mus sp.
XX
PN   WO9921998-A1.
XX
PD   06-MAY-1999.
XX
PF   29-OCT-1998;   98WO-US22991.
XX
PR   14-APR-1998;   98US-0081695.
PR   29-OCT-1997;   97US-0063704.
PR   03-FEB-1998;   98US-0073612.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein DA,  Cohen RL,  Goddard A,  Gurney AL,  Hillan K;
PI   Lawrence DA,  Levine AJ,  Pennica D,  Roy MA,  Wood WI;
XX
DR   WPI; 1999-337420/28.
DR   N-PSDB; AAX76484.
XX
PT   New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS   Claim 9; Page 173-174; 284pp; English.
XX
CC   The present invention describes Wnt-1 induced secreted polypeptides,
CC   WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC   and WISP-3 have homology to connective tissue growth factor (CTGF).
CC   Products from the present invention can be used to treat WISP-related
CC   disorders such as breast, ovarian, and colon cancer or melanoma. The
CC   products can be used to treat arteriosclerosis. The products can also be
CC   used to treat other diseases e.g. benign and malignant tumours,
CC   leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC   blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC   disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC   disorders, bone-related disorders such as osteoporosis, trauma such as
CC   burns, incisions, and other wounds, connective tissue disorders,
CC   catabolic states, testicular-related disorders, and inflammatory,
CC   angiogenic and immunologic disorders including arteriosclerosis. The
CC   products can also be used for detection and diagnosis especially of
CC   individuals with neoplastic cell growth or proliferation. The products
CC   can be used in the production of transgenic or knock-out animals.
CC   Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC   cells.
XX
SQ   Sequence   367 AA;
```

```
  Query Match           86.5%;  Score 1675;  DB 20;  Length 367;
  Best Local Similarity  85.3%;  Pred. No. 2.7e-118;
  Matches  285;  Conservative   20;  Mismatches   29;  Indels    0;  Gaps    0;

Qy      2 FTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAIC 61
          |||||||:|::||:|||||||| ||||||||||||||||||:|||||||||||||||
Db     34 ftpapleetttrpefckwpcecpqspprcplgvslitdgcecckicaqqlgdncteaaic 93

Qy     62 DPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGC 121
          |||||||||||||||||||||||||||||||||| ||:||||||:|||||||| |||
Db     94 dphrglycdysgdrpryaigvcaqvvgvgcvldgvrytngesfqpncrynctcidgtvgc 153

Qy    122 TPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWH 181
          |||||  |||||||  || | :|| |||||||:|||:|||:||  || || | | ||  :
Db    154 tplclsprpprlwcrqprhvrvpgqcceqwvcdddarrprqtalldtrafaasgaveqry 213

Qy    182 RNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKKC 241
          ||||||||||||||:||||:|||||||||:|||||||||||||||||||||   |||||||
Db    214 enciaytspwspcsttcglgistrisnvnarcwpeqesrlcnlrpcdvdiqlhikagkkc 273

Qy    242 LAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSRQ 301
          ||||||| : |||||||:|||:|:|||||||| ||||||||||||| | ||||:| |||||
Db    274 lavyqpeeatnftlagcvstrtyrpkycgvctdnrccipyksktisvdfqcpegpgfsrq 333

Qy    302 VLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
          |||||||||||||||||||||||||||||| ||||
Db    334 vlwinacfcnlscrnpndifadlesypdfeeian 367


RESULT  13
AAR79964
ID   AAR79964 standard; Protein; 349 AA.
XX
AC   AAR79964;
XX
DT   12-JUN-1996  (first entry)
XX
DE   Connective tissue growth factor.
XX
KW   Connective tissue growth factor; CTGF; wound healing; vulnerary;
KW   cell proliferation; cancer; fibrosis; atherosclerosis; diagnosis
KW   therapy; mitogen.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   28
FT                   /label= N-glycosylation_site
FT   Modified-site   225
FT                   /label= N-glycosylation_site
XX
PN   US5408040-A.
XX
PD   18-APR-1995.
XX
PF   30-AUG-1991;   91US-0752427.
XX
PR   30-AUG-1991;   91US-0752427.
PR   14-DEC-1993;   93US-0167628.
XX
PA   (UYSF-) UNIV SOUTH FLORIDA.
XX
PI   Bradham DM,  Grotendorst GR;
XX
DR   WPI; 1995-161147/21.
DR   N-PSDB; AAT04226.
XX
PT   New connective tissue growth factor - used to develop prods. for
PT   wound healing and for diagnosis and therapy of cell proliferative
PT   disorders.
XX
PS   Claim 1; Column 19-20; 12pp; English.
XX
CC   Novel human connective tissue growth factor (CTGF) (AAR79964)
CC   is related immunologically and biologically to platelet-derived
CC   growth factor (PDGF), but is the product of a distinct gene.
CC   CTGF is mitogenic and also a chemotactic agent for cells.  It is
CC   produced by endothelial and fibroblastic cells, and probably acts
CC   as a growth factor in wound healing.  Recombinant CTGF can be obtd.
CC   by expression of cDNA clone DB60R32 (AAT04226) in transformed host
CC   cells.  It is used to accelerate wound healing, and to raise
CC   antibodies useful in detecting disorders associated with overgrowth
CC   of cells, such as cancer, fibrotic diseases and atherosclerosis.
XX
SQ   Sequence   349 AA;


  Query Match           44.9%;  Score 869.5;  DB 16;  Length 349;
  Best Local Similarity  45.6%;  Pred. No. 6.9e-58;
  Matches  156;  Conservative   53;  Mismatches   98;  Indels   35;  Gaps    8;

Qy     12 SRP---QFCKWPCECPPSP-PRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
          |||   | |  || ||  | |||| ||||: ||| ||::||:|||: |||   ||||:||
Db     21 srpavgqncsgpcrcpdepaprcpagvslvldgcgccrvcakqlgelcterdpcdphkgl 80

Qy     68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
          :||: |      ||||    |  |:  |  | :|:||| :||| |||:|||||| ||| :
Db     81 fcdf-gspanrkigvctakdgapcifggtvyrsgesfqssckyqctcldgavgcmplcsm 139

Qy    127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
           || |   || |||| :|| |||:|||::         |:|      ||   | :|
Db    140 dvrlpspdcpfprrvklpgkcceewvcde---------pkdq---tvvgpalaayrledt 187

Qy    183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                    ||:  |: || || :||:|:|||::| || |  |::|||| :|||: |:
Db    188 fgpdptmiranclvqttewsacsktcgmgistrvtndnascrlekqsrlcmvrpceadle 247

Qy    233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
            || ||||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db    248 enikkgkkcirtpkiskpikfelsgctsmktyrakfcgvctdgrcctphrtttlpvefkc 307

Qy    293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
          |||    : :::|  | |: :|   |||| |   : | | :
Db    308 pdgevmkknmmfiktcachyncpgdndifeslyyrkmygdma 349


RESULT  14
AAW09089
ID   AAW09089 standard; Protein; 349 AA.
XX
AC   AAW09089;
XX
DT   26-APR-1997  (first entry)
XX
DE   Human connective tissue growth factor.
XX
KW   Connective tissue growth factor; CTGF; mitogen; cell proliferation;
KW   wound healing; cancer; tumour; fibrosis; glaucoma; atherosclerosis;
KW   scleroderma; arthritis, cirrhosis; scar; diagnosis; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   28
FT                   /label= Glycosylation
FT                   /note= "potential N-glycosylation site"
FT   Modified-site   225
FT                   /label= Glycosylation
FT                   /note= "potential N-glycosylation site"
XX
PN   WO9638172-A1.
XX
PD   05-DEC-1996.
XX
PF   31-MAY-1996;   96WO-US08140.
XX
```

```
PR   31-MAY-1996;   96WO-US08140.
XX
PA   (UYSF-) UNIV SOUTH FLORIDA.
XX
PI   Bradham DM,  Grotendorst GR;
XX
DR   WPI; 1997-042659/04.
DR   N-PSDB; AAT45360;
DR   N-PSDB; AAT58534.
XX
PT   Connective tissue growth factor coding sequence and protein  - used
PT   in the treatment of proliferative disorders and to accelerate wound
PT   healing
XX
PS   Claim 19; Page 50-52; 76pp; English.
XX
CC   Novel human connective tissue growth factor (CTGF) (AAW09089) is a
CC   PDGF-immunorelated protein that may play a significant role in the
CC   normal development, growth and repair of human tissue and probably
CC   functions as a growth factor in wound healing.  CTGF may be involved
CC   in diseases in which there is an overgrowth of connective tissue
CC   cells, such as cancer, tumour formation and growth, fibrotic
CC   diseases (e.g. pulmonary fibrosis, kidney fibrosis, glaucoma) and
CC   atherosclerosis.  Recombinant CTGF can be produced in transformed
CC   host cells utilising a cDNA clone isolated from a HUVEC library.
CC   It can be used to accelerate wound healing.  CTGF inhibitors can be
CC   used to treat atherosclerosis and fibrotic diseases such as
CC   scleroderma, arthritis, liver cirrhosis, and scarring.
XX
SQ   Sequence   349 AA;


  Query Match             44.9%;  Score 869.5;  DB 18;  Length 349;
  Best Local Similarity   45.6%;  Pred. No. 6.9e-58;
  Matches  156;  Conservative   53;  Mismatches   98;  Indels   35;  Gaps    8;

Qy      12 SRP---QFCKWPCECPPSP-PRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
           |||   | |  || ||  | |||| |||| ||| ||::||:|||: |||   ||||:||
Db      21 srpavgqncsgpcrcpdepaprcpagvslvldgcgccrvcakqlgelcterdpcdphkgl 80

Qy      68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
           :||: |      ||||    |  |:  |  | :|:||| :||| |||:|||||| ||| :
Db      81 fcdf-gspanrkigvctakdgapcifggtvyrsgesfqssckyqctcldgavgcmplcsm 139

Qy     127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
            || |   || |||| :|| |||:|||::         |:|      ||   | :|
Db     140 dvrlpspdcpfprrvklpgkcceewvcde---------pkdq---tvvgpalaayrledt 187

Qy     183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                    ||:  |: || || :||:|:|||::| || |  |::|||| :|||: |:
Db     188 fgpdptmiranclvqttewsacsktcgmgistrvtndnascrlekqsrlcmvrpceadle 247

Qy     233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
           || ||||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db     248 enikkgkkcirtpkiskpikfelsgctsmktyrakfcgvctdgrcctphrtttlpvefkc 307

Qy     293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
           |||    : :::|  | |: :|   ||||  |   : | | :
Db     308 pdgevmkknmmfiktcachyncpgdndifeslyyrkmygdma 349


RESULT  15
AAW11302
ID   AAW11302 standard; Protein; 349 AA.
XX
AC   AAW11302;
XX
DT   18-MAR-1997  (first entry)
XX
DE   Connective tissue growth factor.
XX
KW   Connective tissue growth factor; CTGF; human; connective tissue cell;
KW   proliferative disease; platelet-derived growth factor; PDGF; development;
KW   tissue growth; repair; umbilical vein endothelial cell; HUVE cell;
KW   antibody; wound healing; cancer; fibrotic disease; atherosclerosis;
KW   inhibitor; protease degradation; growth factor; therapy.
XX
OS   Homo sapiens.
XX
PN   US5585270-A.
XX
PD   17-DEC-1996.
XX
PF   30-AUG-1991;   91US-0752427.
XX
PR   30-AUG-1991;   91US-0752427.
PR   14-DEC-1993;   93US-0167628.
PR   10-FEB-1995;   95US-0386680.
XX
PA   (UYSF-) UNIV SOUTH FLORIDA.
XX
PI   Bradham DM,  Grotendorst GR;
XX
DR   WPI; 1997-051180/05.
DR   N-PSDB; AAT51234.
XX
PT   New nucleic acid encoding connective tissue growth factor - useful
PT   for accelerating wound healing, also for diagnosis and treatment of
PT   proliferative disease
XX
PS   Claim 9; Column 15-18; 11pp; English.
XX
CC   This sequence represents the human connective tissue growth factor
CC   (CTGF).  CTGF is related immunologically and biologically to
CC   platelet-derived growth factor (PDGF), but is encoded by an unrelated
CC   gene.  CTGF is thought to play a significant role in the normal
CC   development, growth, and repair of human tissue, similarly to PDGF.  The
CC   cDNA encoding this sequence was isolated by screening a cDNA library from
CC   human umbilical vein endothelial (HUVE) cells with anti-PDGF antibodies.
CC   CTGF can be used to accelerate wound healing.  Also, elevated levels of
CC   CTGF may be diagnostic of proliferative diseases involving outgrowth of
CC   connective tissue cells, such as cancer, fibrotic disease and
CC   atherosclerosis.  All of these diseases can be treated with reagents
CC   reactive with CTGF, such as antibodies (which can also serve as assay
CC   reagents).  Antisense nucleic acids, and ribozymes could also be used to
CC   inhibit CTGF production.  The advantage with using CTGF is that it is
CC   more stable, and less susceptible to protease degradation than PDGF, and
CC   other growth factors involved in wound healing.  This is believed to be
CC   due to the high Cys content.
XX
SQ   Sequence   349 AA;


  Query Match             44.9%;  Score 869.5;  DB 18;  Length 349;
  Best Local Similarity   45.6%;  Pred. No. 6.9e-58;
  Matches  156;  Conservative   53;  Mismatches   98;  Indels   35;  Gaps    8;

Qy      12 SRP---QFCKWPCECPPSP-PRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
           |||   | |  || ||  | |||| |||| ||| ||::||:|||: |||   ||||:||
Db      21 srpavgqncsgpcrcpdepaprcpagvslvldgcgccrvcakqlgelcterdpcdphkgl 80

Qy      68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
           :||: |      ||||    |  |:  |  | :|:||| :||| |||:|||||| ||| :
Db      81 fcdf-gspanrkigvctakdgapcifggtvyrsgesfqssckyqctcldgavgcmplcsm 139

Qy     127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
            || |   || |||| :|| |||:|||::         |:|      ||   | :|
Db     140 dvrlpspdcpfprrvklpgkcceewvcde---------pkdq---tvvgpalaayrledt 187

Qy     183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                    ||:  |: || || :||:|:|||::| || |  |::|||| :|||: |:
Db     188 fgpdptmiranclvqttewsacsktcgmgistrvtndnascrlekqsrlcmvrpceadle 247

Qy     233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
           || ||||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db     248 enikkgkkcirtpkiskpikfelsgctsmktyrakfcgvctdgrcctphrtttlpvefkc 307
```

```
Qy     293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
           |||    : :::|  | |: :|   ||||  |   : | | :
Db     308 pdgevmkknmmfiktcachyncpgdndifeslyyrkmygdma 349
```

Search completed: July 26, 2001, 08:36:55
Job time: 962 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 26, 2001, 08:20:53 ; Search time 12.27 Seconds
(without alignments)
562.165 Million cell updates/sec

Title: US-09-325-019-2
Perfect score: 1937
Sequence: 1 DFTPAPLEDTSSRPQFCKWP.....NPNDIFADLESYPDFSEIAN 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	869.5	44.9	349	1	US-08-167-628-2	Sequence 2, Appli
2	869.5	44.9	349	1	US-08-386-680-2	Sequence 2, Appli
3	869.5	44.9	349	1	US-08-459-717-2	Sequence 2, Appli
4	869.5	44.9	349	1	US-08-712-302-2	Sequence 2, Appli
5	869.5	44.9	349	2	US-08-880-031-2	Sequence 2, Appli
6	869.5	44.9	349	3	US-09-054-368-2	Sequence 2, Appli
7	869.5	44.9	349	4	US-09-097-179-2	Sequence 2, Appli
8	869.5	44.9	349	4	US-09-054-274-2	Sequence 2, Appli
9	869.5	44.9	349	4	US-09-080-715-2	Sequence 2, Appli
10	869.5	44.9	349	4	US-09-056-704-2	Sequence 2, Appli
11	869.5	44.9	349	5	PCT-US96-08140-2	Sequence 2, Appli
12	860	44.4	348	1	US-08-468-847B-14	Sequence 14, Appl
13	831.5	42.9	348	1	US-08-468-847B-15	Sequence 15, Appl
14	772.5	39.9	379	1	US-08-468-847B-11	Sequence 11, Appl
15	758	39.1	375	1	US-08-468-847B-13	Sequence 13, Appl
16	743.5	38.4	351	1	US-08-468-847B-16	Sequence 16, Appl
17	726.5	37.5	357	1	US-08-468-847B-17	Sequence 17, Appl
18	705.5	36.4	374	1	US-08-468-847B-12	Sequence 12, Appl
19	705.5	36.4	375	2	US-08-459-101A-2	Sequence 2, Appli
20	239	12.3	93	1	US-08-464-339A-13	Sequence 13, Appl
21	164	8.5	90	1	US-08-464-339A-11	Sequence 11, Appl
22	162	8.4	90	1	US-08-464-339A-12	Sequence 12, Appl
23	150	7.7	2050	2	US-08-347-594A-2	Sequence 2, Appli
24	149	7.7	2813	3	US-08-896-449A-2	Sequence 2, Appli
25	149	7.7	2813	3	US-09-132-652-2	Sequence 2, Appli
26	144	7.4	2556	1	US-08-185-432-17	Sequence 17, Appl
27	142.5	7.4	1170	1	US-08-313-288B-20	Sequence 20, Appl
28	139	7.2	1065	2	US-08-400-159-8	Sequence 8, Appli
29	139	7.2	1148	4	US-08-882-046-4	Sequence 4, Appli
30	139	7.2	1257	3	US-08-611-729A-8	Sequence 8, Appli
31	136	7.0	239	5	PCT-US93-01652-1	Sequence 1, Appli
32	134	6.9	441	3	US-08-985-526-3	Sequence 3, Appli
33	133.5	6.9	2556	1	US-08-083-590A-20	Sequence 20, Appl
34	133.5	6.9	2556	3	US-08-532-384-20	Sequence 20, Appl
35	133	6.9	1219	4	US-08-882-046-5	Sequence 5, Appli
36	133	6.9	2523	1	US-08-185-432-18	Sequence 18, Appl
37	132.5	6.8	589	1	US-07-668-648-2	Sequence 2, Appli
38	132.5	6.8	589	2	US-08-429-998-2	Sequence 2, Appli
39	132.5	6.8	589	2	US-08-431-333-2	Sequence 2, Appli
40	132.5	6.8	589	5	PCT-US91-02321-2	Sequence 2, Appli
41	132	6.8	954	2	US-08-749-169A-3	Sequence 3, Appli
42	132	6.8	954	2	US-09-130-032A-3	Sequence 3, Appli
43	130.5	6.7	855	2	US-08-938-365-3	Sequence 3, Appli
44	130.5	6.7	867	2	US-08-938-365-2	Sequence 2, Appli
45	129.5	6.7	1404	2	US-08-400-159-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-08-167-628-2
; Sequence 2, Application US/08167628
; Patent No. 5408040
;  GENERAL INFORMATION:
;    APPLICANT:  Grotendorst, Gary R.
;    APPLICANT:  Bradham Jr., Douglas M.,
;    TITLE OF INVENTION:  CONNECTIVE TISSUE GROWTH FACTOR
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Spensley Horn Jubas & Lubitz
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/167,628
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/752,427
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wetherell, Jr. Ph.D., John W.
;      REGISTRATION NUMBER:  31,678
;      REFERENCE/DOCKET NUMBER:  PD-1294
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-455-5100
;      TELEFAX:  619-455-5110
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  349 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-167-628-2


  Query Match             44.9%;  Score 869.5;  DB 1;  Length 349;
  Best Local Similarity   45.6%;  Pred. No. 2.2e-73;
  Matches  156;  Conservative   53;  Mismatches   98;  Indels   35;  Gaps    8;

Qy      12 SRP---QFCKWPCECPPSP-PRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
           |||   | |  || ||  | |||| ||||: ||| ||::||:|||: |||   ||||:||
```

```
Db      21 SRPAVGQNCSGPCRCPDEPAPRCPAGVSLVLDGCGCCRVCAKQLGELCTERDPCDPHKGL 80

Qy      68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126

Db      81 FCDF-GSPANRKIGVCTAKDGAPCIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSM 139

Qy     127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182

Db     140 DVRLPSPDCPFPRRVKLPGKCCEEWVCDE---------PKDQ---TVVGPALAAYRLEDT 187

Qy     183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232

Db     188 FGPDPTMIRANCLVQTTEWSACSKTCGMGISTRVTNDNASCRLEKQSRLCMVRPCEADLE 247

Qy     233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292

Db     248 ENIKKGKKCIRTPKISKPIKFELSGCTSMKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKC 307

Qy     293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331

Db     308 PDGEVMKKNMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 349


RESULT   2
US-08-386-680-2
; Sequence 2, Application US/08386680
; Patent No. 5585270
;  GENERAL INFORMATION:
;    APPLICANT:  Grotendorst, Gary R.
;    APPLICANT:  Bradham Jr., Douglas M.,
;    TITLE OF INVENTION:  CONNECTIVE TISSUE GROWTH FACTOR
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Spensley Horn Jubas & Lubitz
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/386,680
;      FILING DATE:  10-FEB-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/167,628
;      FILING DATE:
;      APPLICATION NUMBER:  US/07/752,427
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wetherell, Jr. Ph.D., John W.
;      REGISTRATION NUMBER:  31,678
;      REFERENCE/DOCKET NUMBER:  PD-1294
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-455-5100
;      TELEFAX:  619-455-5110
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  349 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-386-680-2


  Query Match           44.9%;  Score 869.5;  DB 1;  Length 349;
  Best Local Similarity  45.6%;  Pred. No. 2.2e-73;
  Matches  156;  Conservative   53;  Mismatches   98;  Indels   35;  Gaps    8;
```

```
Qy .    12 SRP---QFCKWPCECPPSP-PRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67

Db      21 SRPAVGQNCSGPCRCPDEPAPRCPAGVSLVLDGCGCCRVCAKQLGELCTERDPCDPHKGL 80

Qy      68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126

Db      81 FCDF-GSPANRKIGVCTAKDGAPCIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSM 139

Qy     127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182

Db     140 DVRLPSPDCPFPRRVKLPGKCCEEWVCDE---------PKDQ---TVVGPALAAYRLEDT 187

Qy     183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232

Db     188 FGPDPTMIRANCLVQTTEWSACSKTCGMGISTRVTNDNASCRLEKQSRLCMVRPCEADLE 247

Qy     233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292

Db     248 ENIKKGKKCIRTPKISKPIKFELSGCTSMKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKC 307

Qy     293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331

Db     308 PDGEVMKKNMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 349


RESULT   3
US-08-459-717-2
; Sequence 2, Application US/08459717
; Patent No. 5770209
;  GENERAL INFORMATION:
;    APPLICANT:  Grotendorst, Gary R.
;    APPLICANT:  Bradham Jr., Douglas M.,
;    TITLE OF INVENTION:  CONNECTIVE TISSUE GROWTH FACTOR
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Spensley Horn Jubas & Lubitz
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/459,717
;      FILING DATE:  02-JUN-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/752,427
;      FILING DATE:  30-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wetherell, Jr. Ph.D., John W.
;      REGISTRATION NUMBER:  31,678
;      REFERENCE/DOCKET NUMBER:  PD-1294
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-455-5100
;      TELEFAX:  619-455-5110
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  349 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-459-717-2


  Query Match           44.9%;  Score 869.5;  DB 1;  Length 349;
  Best Local Similarity  45.6%;  Pred. No. 2.2e-73;
  Matches  156;  Conservative   53;  Mismatches   98;  Indels   35;  Gaps    8;
```

```
Qy      12 SRP---QFCKWPCECPPSP-PRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
           |||   | |  || ||  | |||| |||| ||| ||::||:|||: |||   ||||:||
Db      21 SRPAVGQNCSGPCRCPDEPAPRCPAGVSLVLDGCGCCRVCAKQLGELCTERDPCDPHKGL 80

Qy      68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
           :||: |      ||||    |  |:  |  | :|:||| :||| |||:|||||| ||| :
Db      81 FCDF-GSPANRKIGVCTAKDGAPCIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSM 139

Qy     127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
            || |   || |||| :|| |||:|||::         |:|      ||   | :|
Db     140 DVRLPSPDCPFPRRVKLPGKCCEEWVCDE---------PKDQ---TVVGPALAAYRLEDT 187

Qy     183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                    ||:  |: || || :||:|:|||::| || |  |::|||| :|||: |:
Db     188 FGPDPTMIRANCLVQTTEWSACSKTCGMGISTRVTNDNASCRLEKQSRLCMVRPCEADLE 247

Qy     233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
           || ||||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db     248 ENIKKGKKCIRTPKISKPIKFELSGCTSMKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKC 307

Qy     293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
           |||    : :::|  | |: :|   |||| |   : | | :
Db     308 PDGEVMKKNMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 349


RESULT   4
US-08-712-302-2
; Sequence 2, Application US/08712302
; Patent No. 5783187
;  GENERAL INFORMATION:
;    APPLICANT:  Grotendorst, Gary R.
;    APPLICANT:  Bradham Jr., Douglas M.,
;    TITLE OF INVENTION:  CONNECTIVE TISSUE GROWTH FACTOR
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Spensley Horn Jubas & Lubitz
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/712,302
;      FILING DATE:  11-SEP-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/386,680
;      FILING DATE:  10-FEB-1995
;      APPLICATION NUMBER:  US/08/167,628
;      FILING DATE:
;      APPLICATION NUMBER:  US/07/752,427
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wetherell, Jr. Ph.D., John W.
;      REGISTRATION NUMBER:  31,678
;      REFERENCE/DOCKET NUMBER:  PD-1294
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-455-5100
;      TELEFAX:  619-455-5110
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  349 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-712-302-2
```

```
  Query Match           44.9%;  Score 869.5;  DB 1;  Length 349;
  Best Local Similarity  45.6%;  Pred. No. 2.2e-73;
  Matches  156;  Conservative   53;  Mismatches   98;  Indels   35;  Gaps    8;

Qy      12 SRP---QFCKWPCECPPSP-PRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
           |||   | |  || ||  | |||| |||| ||| ||::||:|||: |||   ||||:||
Db      21 SRPAVGQNCSGPCRCPDEPAPRCPAGVSLVLDGCGCCRVCAKQLGELCTERDPCDPHKGL 80

Qy      68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
           :||: |      ||||    |  |:  |  | :|:||| :||| |||:|||||| ||| :
Db      81 FCDF-GSPANRKIGVCTAKDGAPCIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSM 139

Qy     127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
            || |   || |||| :|| |||:|||::         |:|      ||   | :|
Db     140 DVRLPSPDCPFPRRVKLPGKCCEEWVCDE---------PKDQ---TVVGPALAAYRLEDT 187

Qy     183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                    ||:  |: || || :||:|:|||::| || |  |::|||| :|||: |:
Db     188 FGPDPTMIRANCLVQTTEWSACSKTCGMGISTRVTNDNASCRLEKQSRLCMVRPCEADLE 247

Qy     233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
           || ||||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db     248 ENIKKGKKCIRTPKISKPIKFELSGCTSMKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKC 307

Qy     293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
           |||    : :::|  | |: :|   |||| |   : | | :
Db     308 PDGEVMKKNMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 349


RESULT   5
US-08-880-031-2
; Sequence 2, Application US/08880031
; Patent No. 5916756
;  GENERAL INFORMATION:
;    APPLICANT:  Grotendorst, Gary R.
;    APPLICANT:  Bradham Jr., Douglas M.,
;    TITLE OF INVENTION:  CONNECTIVE TISSUE GROWTH FACTOR
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Spensley Horn Jubas & Lubitz
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/880,031
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/167,628
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wetherell, Jr. Ph.D., John W.
;      REGISTRATION NUMBER:  31,678
;      REFERENCE/DOCKET NUMBER:  PD-1294
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-455-5100
;      TELEFAX:  619-455-5110
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  349 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-880-031-2
```

```
  Query Match           44.9%;  Score 869.5;  DB 2;  Length 349;
  Best Local Similarity  45.6%;  Pred. No. 2.2e-73;
  Matches  156;  Conservative   53;  Mismatches   98;  Indels   35;  Gaps    8;

Qy      12 SRP---QFCKWPCECPPSP-PRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
           |||   | |  || ||  | |||| |||| ||| ||::||:|||: |||   ||||:||
Db      21 SRPAVGQNCSGPCRCPDEPAPRCPAGVSLVLDGCGCCRVCAKQLGELCTERDPCDPHKGL 80

Qy      68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
           :||: |      ||||    |  |:  |  | :|:||| :||| |||:|||||| ||| :
Db      81 FCDF-GSPANRKIGVCTAKDGAPCIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSM 139

Qy     127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
           || |   || |||| :|| |||:|||::          |:|      ||   | :|
Db     140 DVRLPSPDCPFPRRVKLPGKCCEEWVCDE---------PKDQ---TVVGPALAAYRLEDT 187

Qy     183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                    ||:  |: || || :||:|:|||::| || |  |::|||| :|||: |:
Db     188 FGPDPTMIRANCLVQTTEWSACSKTCGMGISTRVTNDNASCRLEKQSRLCMVRPCEADLE 247

Qy     233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
           || ||||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db     248 ENIKKGKKCIRTPKISKPIKFELSGCTSMKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKC 307

Qy     293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
           |||    : :::|  | |: :|   ||||  |   : | | :
Db     308 PDGEVMKKNMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 349


RESULT   6
US-09-054-368-2
; Sequence 2, Application US/09054368
; Patent No.. 6069006
; GENERAL INFORMATION:
;  APPLICANT: University of South Florida
;  APPLICANT: Grotendorst, Gary R.
;  APPLICANT: Bradham, Jr., Douglass M.
;  TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
;  FILE REFERENCE: 07414/003005
;  CURRENT APPLICATION NUMBER: US/09/054,368
;  CURRENT FILING DATE: 1998-04-02
;  EARLIER APPLICATION NUMBER: 08/386,680
;  EARLIER FILING DATE: 1995-02-10
;  EARLIER APPLICATION NUMBER: 08/459,717
;  EARLIER FILING DATE: 1995-06-02
;  EARLIER APPLICATION NUMBER: 08/167,628
;  EARLIER FILING DATE: 1993-12-14
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 349
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-09-054-368-2


  Query Match           44.9%;  Score 869.5;  DB 3;  Length 349;
  Best Local Similarity  45.6%;  Pred. No. 2.2e-73;
  Matches  156;  Conservative   53;  Mismatches   98;  Indels   35;  Gaps    8;

Qy      12 SRP---QFCKWPCECPPSP-PRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
           |||   | |  || ||  | |||| |||| ||| ||::||:|||: |||   ||||:||
Db      21 SRPAVGQNCSGPCRCPDEPAPRCPAGVSLVLDGCGCCRVCAKQLGELCTERDPCDPHKGL 80

Qy      68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
           :||: |      ||||    |  |:  |  | :|:||| :||| |||:|||||| ||| :
Db      81 FCDF-GSPANRKIGVCTAKDGAPCIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSM 139

Qy     127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
           || |   || |||| :|| |||:|||::          |:|      ||   | :|
Db     140 DVRLPSPDCPFPRRVKLPGKCCEEWVCDE---------PKDQ---TVVGPALAAYRLEDT 187

Qy     183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                    ||:  |: || || :||:|:|||::| || |  |::|||| :|||: |:
Db     188 FGPDPTMIRANCLVQTTEWSACSKTCGMGISTRVTNDNASCRLEKQSRLCMVRPCEADLE 247

Qy     233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
           || ||||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db     248 ENIKKGKKCIRTPKISKPIKFELSGCTSMKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKC 307

Qy     293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
           |||    : :::|  | |: :|   ||||  |   : | | :
Db     308 PDGEVMKKNMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 349


RESULT   7
US-09-097-179-2
; Sequence 2, Application US/09097179
; Patent No. 6149916
;  GENERAL INFORMATION:
;    APPLICANT:  Grotendorst, Gary R.
;    APPLICANT:  Bradham Jr., Douglas M.,
;    TITLE OF INVENTION:  CONNECTIVE TISSUE GROWTH FACTOR
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Spensley Horn Jubas & Lubitz
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/097,179
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/386,680
;      FILING DATE:  10-FEB-1995
;      APPLICATION NUMBER:  US/08/167,628
;      FILING DATE:
;      APPLICATION NUMBER:  US/07/752,427
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wetherell, Jr. Ph.D., John W.
;      REGISTRATION NUMBER:  31,678
;      REFERENCE/DOCKET NUMBER:  PD-1294
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-455-5100
;      TELEFAX:  619-455-5110
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  349 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-097-179-2


  Query Match           44.9%;  Score 869.5;  DB 4;  Length 349;
  Best Local Similarity  45.6%;  Pred. No. 2.2e-73;
  Matches  156;  Conservative   53;  Mismatches   98;  Indels   35;  Gaps    8;

Qy      12 SRP---QFCKWPCECPPSP-PRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
           |||   | |  || ||  | |||| |||| ||| ||::||:|||: |||   ||||:||
Db      21 SRPAVGQNCSGPCRCPDEPAPRCPAGVSLVLDGCGCCRVCAKQLGELCTERDPCDPHKGL 80

Qy      68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
           :||: |      ||||    |  |:  |  | :|:||| :||| |||:|||||| ||| :
Db      81 FCDF-GSPANRKIGVCTAKDGAPCIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSM 139

Qy     127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
```

```
Db    140 DVRLPSPDCPFPRRVKLPGKCCEEWVCDE---------PKDQ---TVVGPALAAYRLEDT 187

Qy    183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                   ||:  |: || || :||:|:|||::| || |  |::|||| :|||: |:
Db    188 FGPDPTMIRANCLVQTTEWSACSKTCGMGISTRVTNDNASCRLEKQSRLCMVRPCEADLE 247

Qy    233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
           || ||||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db    248 ENIKKGKKCIRTPKISKPIKFELSGCTSMKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKC 307

Qy    293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
          |||    : :::|  | |: :|   ||||  |   : | | :
Db    308 PDGEVMKKNMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 349


RESULT   8
US-09-054-274-2
; Sequence 2, Application US/09054274
; Patent No. 6150101
; GENERAL INFORMATION:
;  APPLICANT: University of South Florida
;  APPLICANT: Grotendorst, Gary R.
;  APPLICANT: Bradham, Jr., Douglass M.
;  TITLE OF INVENTION: METHODS OF IDENTIFYING A COMPOSITION
;  TITLE OF INVENTION: THAT ALTERS CONNECTIVE TISSUE GROWTH
;  TITLE OF INVENTION: FACTOR EXPRESSION (Amended)
;  FILE REFERENCE: 07414/003004
;  CURRENT APPLICATION NUMBER: US/09/054,274
;  CURRENT FILING DATE: 1998-04-02
;  EARLIER APPLICATION NUMBER: 08/386,680
;  EARLIER FILING DATE: 1995-02-10
;  EARLIER APPLICATION NUMBER: 08/459,717
;  EARLIER FILING DATE: 1995-06-02
;  EARLIER APPLICATION NUMBER: 08/167,628
;  EARLIER FILING DATE: 1993-12-14
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 349
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-09-054-274-2


  Query Match             44.9%;  Score 869.5;  DB 4;  Length 349;
  Best Local Similarity   45.6%;  Pred. No. 2.2e-73;
  Matches  156;  Conservative   53;  Mismatches   98;  Indels   35;  Gaps    8;

Qy     12 SRP---QFCKWPCECPPSP-PRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
             |||   | |  || ||  | |||| |||| :||| ||::||:|||: |||   ||||:||
Db     21 SRPAVGQNCSGPCRCPDEPAPRCPAGVSLVLDGCGCCRVCAKQLGELCTERDPCDPHKGL 80

Qy     68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
          :||: |      ||||    |  |:  |  | :|:||| :||| |||:|||||| ||| :
Db     81 FCDF-GSPANRKIGVCTAKDGAPCIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSM 139

Qy    127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
           || |   || |||| :|| |||:|||::          |:|      ||   | :|
Db    140 DVRLPSPDCPFPRRVKLPGKCCEEWVCDE---------PKDQ---TVVGPALAAYRLEDT 187

Qy    183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                   ||:  |: || || :||:|:|||::| || |  |::|||| :|||: |:
Db    188 FGPDPTMIRANCLVQTTEWSACSKTCGMGISTRVTNDNASCRLEKQSRLCMVRPCEADLE 247

Qy    233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
           || ||||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db    248 ENIKKGKKCIRTPKISKPIKFELSGCTSMKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKC 307

Qy    293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
          |||    : :::|  | |: :|   ||||  |   : | | :
Db    308 PDGEVMKKNMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 349


RESULT   9
US-09-080-715-2
; Sequence 2, Application US/09080715
; Patent No. 6190884
;  GENERAL INFORMATION:
;    APPLICANT:  Grotendorst, Gary R.
;    APPLICANT:  Bradham Jr., Douglas M.,
;    TITLE OF INVENTION:  CONNECTIVE TISSUE GROWTH FACTOR
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Spensley Horn Jubas & Lubitz
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/080,715
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/167,628
;      FILING DATE:
;      APPLICATION NUMBER:  US/07/752,427
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wetherell, Jr. Ph.D., John W.
;      REGISTRATION NUMBER:  31,678
;      REFERENCE/DOCKET NUMBER:  PD-1294
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-455-5100
;      TELEFAX:  619-455-5110
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  349 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-080-715-2


  Query Match             44.9%;  Score 869.5;  DB 4;  Length 349;
  Best Local Similarity   45.6%;  Pred. No. 2.2e-73;
  Matches  156;  Conservative   53;  Mismatches   98;  Indels   35;  Gaps    8;

Qy     12 SRP---QFCKWPCECPPSP-PRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
             |||   | |  || ||  | |||| |||| :||| ||::||:|||: |||   ||||:||
Db     21 SRPAVGQNCSGPCRCPDEPAPRCPAGVSLVLDGCGCCRVCAKQLGELCTERDPCDPHKGL 80

Qy     68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
          :||: |      ||||    |  |:  |  | :|:||| :||| |||:|||||| ||| :
Db     81 FCDF-GSPANRKIGVCTAKDGAPCIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSM 139

Qy    127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
           || |   || |||| :|| |||:|||::          |:|      ||   | :|
Db    140 DVRLPSPDCPFPRRVKLPGKCCEEWVCDE---------PKDQ---TVVGPALAAYRLEDT 187

Qy    183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                   ||:  |: || || :||:|:|||::| || |  |::|||| :|||: |:
Db    188 FGPDPTMIRANCLVQTTEWSACSKTCGMGISTRVTNDNASCRLEKQSRLCMVRPCEADLE 247

Qy    233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
           || ||||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db    248 ENIKKGKKCIRTPKISKPIKFELSGCTSMKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKC 307

Qy    293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
          |||    : :::|  | |: :|   ||||  |   : | | :
```

```
Db    308 PDGEVMKKNMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 349


RESULT  10
US-09-056-704-2
; Sequence 2, Application US/09056704
; Patent No. 6232064
; GENERAL INFORMATION:
;  APPLICANT: University of South Florida
;  APPLICANT: Grotendorst, Gary R.
;  APPLICANT: Bradham, Jr., Douglass M.
;  TITLE OF INVENTION: METHODS OF DIAGNOSING A PATHOLOGY
;  TITLE OF INVENTION: CHARACTERIZED BY A CELL PROLIFERATIVE
;  TITLE OF INVENTION: DISORDER ASSOCIATED WITH CONNECTIVE
;  TITLE OF INVENTION: TISSUE GROWTH FACTOR (Amended)
;  FILE REFERENCE: 07414/003002
;  CURRENT APPLICATION NUMBER: US/09/056,704
;  CURRENT FILING DATE: 1998-03-03
;  EARLIER APPLICATION NUMBER: 08/386,680
;  EARLIER FILING DATE: 1995-02-10
;  EARLIER APPLICATION NUMBER: 08/459,717
;  EARLIER FILING DATE: 1995-06-02
;  EARLIER APPLICATION NUMBER: 08/167,628
;  EARLIER FILING DATE: 1993-12-14
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 349
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-09-056-704-2


  Query Match             44.9%;  Score 869.5;  DB 4;  Length 349;
  Best Local Similarity   45.6%;  Pred. No. 2.2e-73;
  Matches  156;  Conservative   53;  Mismatches   98;  Indels   35;  Gaps    8;

Qy      12 SRP---QFCKWPCECPPSP-PRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
           |||   | |  || ||  | |||| |||| ||| ||::||:|||: |||   ||||:||
Db      21 SRPAVGQNCSGPCRCPDEPAPRCPAGVSLVLDGCGCCRVCAKQLGELCTERDPCDPHKGL 80

Qy      68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
           :||: |      ||||    |  |:  |  | :|:||| :||| |||:|||||| ||| :
Db      81 FCDF-GSPANRKIGVCTAKDGAPCIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSM 139

Qy     127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
            || |   || |||| :|| |||:|||::         |:|       ||   | :|
Db     140 DVRLPSPDCPFPRRVKLPGKCCEEWVCDE---------PKDQ---TVVGPALAAYRLEDT 187

Qy     183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                     ||:  |: || || :||:|:|||::| || |  |::|||| :|||: |:
Db     188 FGPDPTMIRANCLVQTTEWSACSKTCGMGISTRVTNDNASCRLEKQSRLCMVRPCEADLE 247

Qy     233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
           || ||||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db     248 ENIKKGKKCIRTPKISKPIKFELSGCTSMKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKC 307

Qy     293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
           |||    : :::|  | |: :|   ||||  |   : | | :
Db     308 PDGEVMKKNMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 349


RESULT  11
PCT-US96-08140-2
; Sequence 2, Application PC/TUS9608140
;  GENERAL INFORMATION:
;    APPLICANT:  University of South Florida
;    TITLE OF INVENTION:  CONNECTIVE TISSUE GROWTH FACTOR
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FISH & RICHARDSON P.C.
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US96/08140
;      FILING DATE:  30-MAY-1996
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haile, Ph.D., Lisa A.
;      REGISTRATION NUMBER:  38,347
;      REFERENCE/DOCKET NUMBER:  07414/003WO1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-678-5070
;      TELEFAX:  619-678-5099
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  349 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US96-08140-2


  Query Match             44.9%;  Score 869.5;  DB 5;  Length 349;
  Best Local Similarity   45.6%;  Pred. No. 2.2e-73;
  Matches  156;  Conservative   53;  Mismatches   98;  Indels   35;  Gaps    8;

Qy      12 SRP---QFCKWPCECPPSP-PRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
           |||   | |  || ||  | |||| |||| ||| ||::||:|||: |||   ||||:||
Db      21 SRPAVGQNCSGPCRCPDEPAPRCPAGVSLVLDGCGCCRVCAKQLGELCTERDPCDPHKGL 80

Qy      68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
           :||: |      ||||    |  |:  |  | :|:||| :||| |||:|||||| ||| :
Db      81 FCDF-GSPANRKIGVCTAKDGAPCIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSM 139

Qy     127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
            || |   || |||| :|| |||:|||::         |:|       ||   | :|
Db     140 DVRLPSPDCPFPRRVKLPGKCCEEWVCDE---------PKDQ---TVVGPALAAYRLEDT 187

Qy     183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                     ||:  |: || || :||:|:|||::| || |  |::|||| :|||: |:
Db     188 FGPDPTMIRANCLVQTTEWSACSKTCGMGISTRVTNDNASCRLEKQSRLCMVRPCEADLE 247

Qy     233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
           || ||||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db     248 ENIKKGKKCIRTPKISKPIKFELSGCTSMKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKC 307

Qy     293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
           |||    : :::|  | |: :|   ||||  |   : | | :
Db     308 PDGEVMKKNMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 349


RESULT  12
US-08-468-847B-14
; Sequence 14, Application US/08468847B
; Patent No. 5780263
;  GENERAL INFORMATION:
;    APPLICANT:  Hastings, Gregg A. and Adams, Mark D.
;    TITLE OF INVENTION:  Human CCN-Like Growth Factor
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILIAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07068
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/468,847B
;      FILING DATE:  6 June 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MULLINS, J.G.
;      REGISTRATION NUMBER:  33,073
;      REFERENCE/DOCKET NUMBER:  325800-442
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 AMINO ACIDS
;      TYPE:  AMINO ACID
;      STRANDEDNESS:
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  PROTEIN
US-08-468-847B-14


  Query Match             44.4%;  Score 860;  DB 1;  Length 348;
  Best Local Similarity   45.6%;  Pred. No. 1.7e-72;
  Matches  156;  Conservative   53;  Mismatches   97;  Indels   36;  Gaps    9;

Qy      12 SRP---QFCKWPCECPPSP-PRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
           |||   | |  || ||  | |||| ||||: ||| ||::||:|||: |||   ||||:||
Db      21 SRPAVGQNCSGPCRCPDEPAPRCPAGVSLV-DGCGCCRVCAKQLGELCTERDPCDPHKGL 79

Qy      68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
           :||: |      ||||    |  |:  |  | :|:||| :||| |||:|||||| ||| :
Db      80 FCDF-GSPANRKIGVCTAKDGAPCIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSM 138

Qy     127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
            || |   || |||| :|| |||:|||::         |:|      ||   | :|
Db     139 DVRLPSPDCPFPRRVKLPGKCCEEWVCDE---------PKDQ---TVVGPALAAYRLEDT 186

Qy     183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                    ||:  |: || || :||:|:|||::| || |  |::|||| :|||: |:
Db     187 FGPDPTMIRANCLVQTTEWSACSKTCGMGISTRVTNDNASCRLEKQSRLCMVRPCEADLE 246

Qy     233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
           || ||||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db     247 ENIKKGKKCIRTPKISKPIKFELSGCTSMKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKC 306

Qy     293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
           |||    : :::|  | |: :|   ||||  |   : | | :
Db     307 PDGEVMKKNMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 348


RESULT  13
US-08-468-847B-15
; Sequence 15, Application US/08468847B
; Patent No. 5780263
;  GENERAL INFORMATION:
;    APPLICANT:  Hastings, Gregg A. and Adams, Mark D.
;    TITLE OF INVENTION:  Human CCN-Like Growth Factor
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILIAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/468,847B
;      FILING DATE:  6 June 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MULLINS, J.G.
;      REGISTRATION NUMBER:  33,073
;      REFERENCE/DOCKET NUMBER:  325800-442
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 AMINO ACIDS
;      TYPE:  AMINO ACID
;      STRANDEDNESS:
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  PROTEIN
US-08-468-847B-15


  Query Match             42.9%;  Score 831.5;  DB 1;  Length 348;
  Best Local Similarity   44.2%;  Pred. No. 7.7e-70;
  Matches  151;  Conservative   56;  Mismatches  100;  Indels   35;  Gaps    8;

Qy      12 SRP---QFCKWPCEC-PPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
           :||   | |   |:|   : | || ||||: ||| ||::||:|||: |||   ||||:||
Db      20 TRPATGQDCSAQCQCAAEAAPHCPAGVSLVLDGCGCCRVCAKQLGELCTERDPCDPHKGL 79

Qy      68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
           :||: |      ||||    |  ||  |  | :|:||| :||| |||:|||||| ||| :
Db      80 FCDF-GSPANRKIGVCTAKDGAPCVFGGSVYRSGESFQSSCKYQCTCLDGAVGCVPLCSM 138

Qy     127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
            || |   || |||| :|| ||::|||::         |:|     |||   | :|
Db     139 DVRLPSPDCPFPRRVKLPGKCCKEWVCDE---------PKDR---TAVGPALAAYRLEDT 186

Qy     183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                    ||:  |: || || :||:|:|||::| |  |  |::|||| :|||: |:
Db     187 FGPDPTMMRANCLVQTTEWSACSKTCGMGISTRVTNDNTFCRLEKQSRLCMVRPCEADLE 246

Qy     233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
           || ||||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db     247 ENIKKGKKCIRTPKIAKPVKFELSGCTSVKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKC 306

Qy     293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
           |||    : :::|  | |: :|   ||||  |   : | | :
Db     307 PDGEIMKKNMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 348


RESULT  14
US-08-468-847B-11
; Sequence 11, Application US/08468847B
; Patent No. 5780263
;  GENERAL INFORMATION:
;    APPLICANT:  Hastings, Gregg A. and Adams, Mark D.
;    TITLE OF INVENTION:  Human CCN-Like Growth Factor
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILIAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
```

```
;       COUNTRY:  USA
;       ZIP:  07068
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5 INCH DISKETTE
;       COMPUTER:  IBM PS/2
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  WORD PERFECT 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/468,847B
;       FILING DATE:  6 June 1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  MULLINS, J.G.
;       REGISTRATION NUMBER:  33,073
;       REFERENCE/DOCKET NUMBER:  325800-442
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  201-994-1700
;       TELEFAX:  201-994-1744
;   INFORMATION FOR SEQ ID NO:  11:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  379 AMINO ACIDS
;       TYPE:  AMINO ACID
;       STRANDEDNESS:
;       TOPOLOGY:  LINEAR
;     MOLECULE TYPE:  PROTEIN
US-08-468-847B-11


  Query Match             39.9%;  Score 772.5;  DB 1;  Length 379;
  Best Local Similarity   39.6%;  Pred. No. 2.6e-64;
  Matches  141;  Conservative   56;  Mismatches  114;  Indels   45;  Gaps    6;

Qy      17 CKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGLYCDYSGDRP 76
           |   | ||   |:|  || |: ||| |||:||:|| ::|::   ||  :|| |:: |
Db      26 CPAACHCPLEAPKCAPGVGLVRDGCGCCKVCAKQLNEDCSKTQPCDHTKGLECNF-GASS 84

Qy      77 RYAIGVC-AQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-LRVRPPRLW 134
             |:| ||  |  |  :   | ||:||||||: |||||||||| |||   :  | |
Db      85 TALKGICRAQSEGRPCEYNSRIYQNGESFQPNCKHQCTCIDGAVGCIPLCPQELSLPNLG 144

Qy     135 CPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHRN---------- 183
           ||:|| | : | |||:|||::|: :       |    ||  |||    |
Db     145 CPNPRLVKVSGQCCEEWVCDEDSIKDSLDDQDDLLGLDA-SEVELTRNNELIAIGKGSSL 203

Qy     184 ----------------------CIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESR 220
                                 ||  |: || || ||| |:|||::| | :|   :|:|
Db     204 KRLPVFGTEPRVLFNPLHAHGQKCIVQTTSWSQCSKSCGTGISTRVTNDNPECRLVKETR 263

Qy     221 LCNLRPCDVDIHTLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIP 280
           :| :|||   ::: :| ||||    :    : || ||| | : |:|||| |:| ||| |
Db     264 ICEVRPCGQPVYSSLKKGKKCSKTKKSPEPVRFTYAGCSSVKKYRPKYCGSCVDGRCCTP 323

Qy     281 YKSKTIDVSFQCPDGLGFSRQVLWINACFCNLSCRNPND-------IFADLESYPD 329
           :::|: : |:| ||  ||: |: | :| || :| :|||:       :| |:  : |
Db     324 LQTRTVKMRFRCEDGEMFSKNVMMIQSCKCNYNCPHPNEASFRLYSLFNDIHKFRD 379


RESULT  15
US-08-468-847B-13
; Sequence 13, Application US/08468847B
; Patent No. 5780263
;  GENERAL INFORMATION:
;    APPLICANT:  Hastings, Gregg A. and Adams, Mark D.
;    TITLE OF INVENTION:  Human CCN-Like Growth Factor
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILIAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;       STATE:  NEW JERSEY
;       COUNTRY:  USA
;       ZIP:  07068
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5 INCH DISKETTE
;       COMPUTER:  IBM PS/2
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  WORD PERFECT 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/468,847B
;       FILING DATE:  6 June 1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  MULLINS, J.G.
;       REGISTRATION NUMBER:  33,073
;       REFERENCE/DOCKET NUMBER:  325800-442
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  201-994-1700
;       TELEFAX:  201-994-1744
;   INFORMATION FOR SEQ ID NO:  13:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  375 AMINO ACIDS
;       TYPE:  AMINO ACID
;       STRANDEDNESS:
;       TOPOLOGY:  LINEAR
;     MOLECULE TYPE:  PROTEIN
US-08-468-847B-13


  Query Match             39.1%;  Score 758;  DB 1;  Length 375;
  Best Local Similarity   38.9%;  Pred. No. 5.8e-63;
  Matches  142;  Conservative   55;  Mismatches  100;  Indels   68;  Gaps    9;

Qy      17 CKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGLYCDYSGDRP 76
           |   |:|| : |:|  || |: ||| |||:||:|| ::|:    ||  :|| |:: |  |
Db      26 CPAVCQCPAAAPQCAPGVGLVPDGCGCCKVCAKQLNEDCSRTQPCDHTKGLECNF-GASP 84

Qy      77 RYAIGVC-AQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-LRVRPPRLW 134
             |:| ||  |  |  :   | ||:||||||: ||||||||| |||   :  | |
Db      85 AATNGICRAQSEGRPCEYNSKIYQNGESFQPNCKHQCTCIDGAVGCIPLCPQELSLPNLG 144

Qy     135 CPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWH------------ 181
           || || | :|| |||:||| |::|              ||: |:| :
Db     145 CPSPRLVKVPGQCCEEWVC-DESK------------DALEELEGFFSKEFGLDASEGEL 190

Qy     182 -RN----------------------------CIAYTSPWSPCSTSCGLGVSTRISNVN 210
            ||                            ||  |: || || :|| |:|||::| |
Db     191 TRNNELIAIVKGGLKMLPVFGSEPQSRAFENPKCIVQTTSWSQCSKTCGTGISTRVTNDN 250

Qy     211 AQCWPEQESRLCNLRPCDVDIHTLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCG 270
             |   :|:|:| :|||    :  :| ||||    :  : : || ||| | : |:|||||
Db     251 PDCKLIKETRICEVRPCGQPSYASLKKGKKCTKTKKSPSPVRFTYAGCSSVKKYRPKYCG 310

Qy     271 VCMDNRCCIPYKSKTIDVSFQCPDGLGFSRQVLWINACFCNLSCRNPNDIFADLESYPDF 330
           |:| ||| | :::|: : |:| ||  |:: |: | :| || :| : |      |:|| |
Db     311 SCVDGRCCTPQQTRTVKIRFRCDDGETFTKSVMMIQSCRCNYNCPHAN------EAYP-F 363

Qy     331 SEIAN 335
             : |
Db     364 YRLVN 368


Search completed: July 26, 2001, 08:34:31
Job time: 818 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 26, 2001, 08:20:53 ; Search time 16.04 Seconds
(without alignments)
1590.927 Million cell updates/sec

Title: US-09-325-019-2
Perfect score: 1937
Sequence: 1 DFTPAPLEDTSSRPQFCKWP.....NPNDIFADLESYPDFSEIAN 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	869.5	44.9	349	2	A40551	connective tissue
2	835.5	43.1	348	2	A40578	beta IG-M2 protein
3	772.5	39.9	379	2	A35669	gene CYR61 protein
4	758	39.1	375	2	A41428	CEF-10 protein pre
5	743.5	38.4	351	2	S20078	NOV protein - chic
6	726.5	37.5	357	2	I38069	gene novH protein
7	151.5	7.8	3020	2	A43932	mucin 2 precursor,
8	150	7.7	2531	2	A46019	Notch-1 protein -
9	150	7.7	2813	1	VWHU	von Willebrand fac
10	147.5	7.6	1178	1	A39804	thrombospondin pre
11	145.5	7.5	1042	2	A57534	mucin 5AC (clone L
12	144.5	7.5	837	2	A42112	mucin-like peptide
13	144	7.4	1444	2	T18856	angiogenesis inhib
14	142.5	7.4	1170	1	TSHUP1	thrombospondin 1 p
15	142	7.3	1111	2	T26972	hypothetical prote
16	140.5	7.3	1170	2	A40558	thrombospondin 1 p
17	139.5	7.2	4135	2	T42629	tenascin-X - bovin
18	139.5	7.2	5376	2	T42215	zonadhesin - mouse
19	139	7.2	1700	2	S08167	Balbiani ring 3 pr
20	138.5	7.2	2139	2	A35672	crumbs protein - f
21	138	7.1	1056	2	A53767	mucin MUC5B, trach
22	138	7.1	2555	2	A40043	notch protein homo
23	138	7.1	3566	1	A40701	tenascin-X precurs
24	137.5	7.1	810	2	T10756	Nel-homolog protei
25	137	7.1	4006	2	T09070	probable tenascin
26	136	7.0	2437	2	S42612	transmembrane prot
27	135	7.0	2531	2	S18188	notch protein homo
28	134.5	6.9	854	1	QRHYLD	LDL receptor precu
29	134.5	6.9	1034	2	JC5598	mucin - rat
30	133	6.9	1220	2	A56136	jagged protein pre
31	133	6.9	1360	2	T33922	hypothetical prote
32	133	6.9	1620	2	T27283	hypothetical prote
33	133	6.9	2524	2	A35844	Xotch protein - Af
34	132.5	6.8	589	2	B38128	epithelin/granulin
35	132.5	6.8	862	1	QRMSLD	LDL receptor precu
36	131.5	6.8	1964	2	T09059	notch4 - mouse
37	129.5	6.7	835	2	JP0076	nel protein - chic
38	129.5	6.7	1408	2	S16148	gene serrate prote
39	129.5	6.7	2406	2	A54148	odz protein - frui
40	129.5	6.7	2515	2	S47008	tenascin-like prot
41	129	6.7	1172	1	TSHUP2	thrombospondin 2 p
42	129	6.7	1737	2	T00209	MEGF8 protein - hu
43	128.5	6.6	1203	2	A49175	Motch B protein -
44	128.5	6.6	2918	2	A54105	fibrillin-2 precur
45	128.5	6.6	4544	1	S02392	alpha-2-macroglobu

ALIGNMENTS

RESULT 1
A40551
connective tissue growth factor - human
C;Species: Homo sapiens (man)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 21-Jul-2000
C;Accession: A40551; S44205
R;Bradham, D.M.; Igarashi, A.; Potter, R.L.; Grotendorst, G.R.
J. Cell Biol. 114, 1285-1294, 1991
A;Title: Connective tissue growth factor: a cysteine-rich mitogen secreted by human v
A;Reference number: A40551; MUID:91373462
A;Accession: A40551
A;Molecule type: mRNA
A;Residues: 1-349 <BRA>
A;Cross-references: GB:M92934; GB:M36965; GB:S56201; NID:g180923; PIDN:AAA91279.1; PI
R;Oemar, B.S.; Werner, A.; Yang, Z.; Garnier, J.M.; Gentz, R.; Luescher, T.F.
submitted to the EMBL Data Library, April 1994
A;Description: Differential cloning and expression of human connective tissue growth
A;Reference number: S44205
A;Accession: S44205
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-349 <OEM>
A;Cross-references: EMBL:X78947; NID:g474933; PID:g474934

Query Match 44.9%; Score 869.5; DB 2; Length 349;
Best Local Similarity 45.6%; Pred. No. 2.5e-57;
Matches 156; Conservative 53; Mismatches 98; Indels 35; Gaps 8;

```
Qy      12 SRP---QFCKWPCECPPSP-PRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
           |||   | |  || ||  | |||| |||| ||| ||::||:||| ||| ||||:||
Db      21 SRPAVGQNCSGPCRCPDEPAPRCPAGVSLVLDGCGCCRVCAKQLGELCTERDPCDPHKGL 80

Qy      68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
           :||: |      ||||    |  |:  |  | :|:||| :||| |||:|||||| ||| :
Db      81 FCDF-GSPANRKIGVCTAKDGAPCIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSM 139

Qy     127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
           || |   || |||| :|| |||:|||::         |:|      ||   | :|
Db     140 DVRLPSPDCPFPRRVKLPGKCCEEWVCDE---------PKDQ---TVVGPALAAYRLEDT 187

Qy     183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                    ||:  |: || || :||:|:|||::| || |  |::|||| :|||: |:
Db     188 FGPDPTMIRANCLVQTTEWSACSKTCGMGISTRVTNDNASCRLEKQSRLCMVRPCEADLE 247

Qy     233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
             || ||||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db     248 ENIKKGKKCIRTPKISKPIKFELSGCTSMKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKC 307

Qy     293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
           |||    : :::|  | |: :|   ||||  |   : | | :
Db     308 PDGEVMKKNMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 349
```

```
RESULT   2
A40578
beta IG-M2 protein precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 06-Mar-1992 #sequence_revision 06-Mar-1992 #text_change 01-Dec-2000
C;Accession: A40578; A53228
R;Brunner, A.; Chinn, J.; Neubauer, M.; Purchio, A.F.
DNA Cell Biol. 10, 293-300, 1991
A;Title: Identification of a gene family regulated by transforming growth factor-beta.
A;Reference number: A40578; MUID:91229699
A;Accession: A40578
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-348 <BRU>
A;Cross-references: GB:M80263; NID:g201945; PIDN:AAA73135.1; PID:g201946
R;Ryseck, R.P.; Macdonald-Bravo, H.; Mattei, M.G.; Bravo, R.
Cell Growth Differ. 2, 225-233, 1991
A;Title: Structure, mapping, and expression of fisp-12, a growth factor-inducible gene e
A;Reference number: A53228; MUID:91363290
A;Accession: A53228
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-160,'K',162-348 <RYS>
A;Cross-references: GB:M70641; NID:g193313; PIDN:AAA37627.1; PID:g193314
C;Genetics:
A;Gene: fisp-12


  Query Match             43.1%;  Score 835.5;  DB 2;  Length 348;
  Best Local Similarity   44.4%;  Pred. No. 8.2e-55;
  Matches  152;  Conservative   55;  Mismatches  100;  Indels   35;  Gaps    8;

Qy      12 SRP---QFCKWPCEC-PPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
           :||   | |   |:|   : | || |||||: ||| ||::||:|||: |||   ||||:||
Db      20 TRPATGQDCSAQCQCAAEAAPHCPAGVSLVLDGCGCCRVCAKQLGELCTERDPCDPHKGL 79

Qy      68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
           :||: |      ||||    |  ||  |  | :|:||| :||| |||:|||||| ||| :
Db      80 FCDF-GSPANRKIGVCTAKDGAPCVFGGSVYRSGESFQSSCKYQCTCLDGAVGCVPLCSM 138

Qy     127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
            || |   || |||| :|| |||:|||::         |:|     |||   | :|
Db     139 DVRLPSPDCPFPRRVKLPGKCCEEWVCDE---------PKDR---TAVGPALAAYRLEDT 186

Qy     183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                     ||:  |: || || :||:|:|||::| |  |  |::|||| :|||: |:
Db     187 FGPDPTMMRANCLVQTTEWSACSKTCGMGISTRVTNDNTFCRLEKQSRLCMVRPCEADLE 246

Qy     233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
             || ||||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db     247 ENIKKGKKCIRTPKIAKPVKFELSGCTSVKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKC 306

Qy     293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
           |||    : :::|  | |: :|   ||||  |   : | | :
Db     307 PDGEIMKKNMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 348


RESULT   3
A35669
gene CYR61 protein precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 28-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 05-Nov-1999
C;Accession: A35669; I48319; S16446
R;O'Brien, T.P.; Yang, G.P.; Sanders, L.; Lau, L.F.
Mol. Cell. Biol. 10, 3569-3577, 1990
A;Title: Expression of cyr61, a growth factor-inducible immediate-early gene.
A;Reference number: A35669; MUID:90287146
A;Accession: A35669
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-379 <OAB>
A;Cross-references: GB:M32490; NID:g192909; PIDN:AAA37512.1; PID:g309206
A;Note: the authors translated the codon GAT for residue 337 as Gln
R;Latinkic, B.V.; O'Brien, T.P.; Lau, L.F.
Nucleic Acids Res. 19, 3261-3267, 1991
A;Title: Promoter function and structure of the growth factor-inducible immediate ear
A;Reference number: I48319; MUID:91288203
A;Accession: I48319
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-379 <RES>
A;Cross-references: EMBL:X56790; NID:g50632; PIDN:CAA40109.1; PID:g50633
A;Note: the authors did not translate the codon for residue 108
A;Note: the authors translated the codon GAT for residue 337 as Gln
C;Genetics:
A;Gene: CYR61
A;Introns: 21/3; 93/1; 208/1; 279/3
C;Superfamily: von Willebrand factor type C repeat homology
F;99-166/Domain: von Willebrand factor type C repeat homology <VWC>


  Query Match             39.9%;  Score 772.5;  DB 2;  Length 379;
  Best Local Similarity   39.6%;  Pred. No. 4e-50;
  Matches  141;  Conservative   56;  Mismatches  114;  Indels   45;  Gaps    6;

Qy      17 CKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGLYCDYSGDRP 76
           |   | ||   |:|  || |: ||| |||:||:|| ::|::   ||  :|| |:: |
Db      26 CPAACHCPLEAPKCAPGVGLVRDGCGCCKVCAKQLNEDCSKTQPCDHTKGLECNF-GASS 84

Qy      77 RYAIGVC-AQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-LRVRPPRLW 134
             |:| ||  |  |  :   | ||:|||||||: |||||||||| |||   :  | |
Db      85 TALKGICRAQSEGRPCEYNSRIYQNGESFQPNCKHQCTCIDGAVGCIPLCPQELSLPNLG 144

Qy     135 CPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHRN---------- 183
           ||:|| | : | |||:|||::|: :       |    ||  |||    |
Db     145 CPNPRLVKVSGQCCEEWVCDEDSIKDSLDDQDDLLGLDA-SEVELTRNNELIAIGKGSSL 203

Qy     184 ----------------------CIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESR 220
                                 ||  |: || || ||| |:|||::| | :|   :|:|
Db     204 KRLPVFGTEPRVLFNPLHAHGQKCIVQTTSWSQCSKSCGTGISTRVTNDNPECRLVKETR 263

Qy     221 LCNLRPCDVDIHTLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIP 280
           :| :|||   ::: :| ||||    :    : || ||| | : |:|||| |:| ||| |
Db     264 ICEVRPCGQPVYSSLKKGKKCSKTKKSPEPVRFTYAGCSSVKKYRPKYCGSCVDGRCCTP 323

Qy     281 YKSKTIDVSFQCPDGLGFSRQVLWINACFCNLSCRNPND-------IFADLESYPD 329
           :::|: : |:| ||  ||: |: | :| || :| :|||:       :| |:  : |
Db     324 LQTRTVKMRFRCEDGEMFSKNVMMIQSCKCNYNCPHPNEASFRLYSLFNDIHKFRD 379


RESULT   4
A41428
CEF-10 protein precursor - chicken
C;Species: Gallus gallus (chicken)
C;Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 21-Jul-2000
C;Accession: A41428
R;Simmons, D.L.; Levy, D.B.; Yannoni, Y.; Erikson, R.L.
Proc. Natl. Acad. Sci. U.S.A. 86, 1178-1182, 1989
A;Title: Identification of a phorbol ester-repressible v-src-inducible gene.
A;Reference number: A41428; MUID:89145206
A;Accession: A41428
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-375 <SIM>
A;Cross-references: GB:J04496; NID:g211435; PIDN:AAA48661.1; PID:g211436


  Query Match             39.1%;  Score 758;  DB 2;  Length 375;
  Best Local Similarity   38.9%;  Pred. No. 4.7e-49;
  Matches  142;  Conservative   55;  Mismatches  100;  Indels   68;  Gaps    9;

Qy      17 CKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGLYCDYSGDRP 76
           |   |:|| : |:|  || |: ||| |||:||:|| ::|:    ||  :|| |:: |  |
Db      26 CPAVCQCPAAAPQCAPGVGLVPDGCGCCKVCAKQLNEDCSRTQPCDHTKGLECNF-GASP 84
```

```
Qy      77 RYAIGVC-AQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-LRVRPPRLW 134
           |:| ||  |  |  :   | ||:||||||: ||||||||| |||   :  | |
Db      85 AATNGICRAQSEGRPCEYNSKIYQNGESFQPNCKHQCTCIDGAVGCIPLCPQELSLPNLG 144

Qy     135 CPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWH------------ 181
           || || | :|| |||:||| |::|             ||: |:| :
Db     145 CPSPRLVKVPGQCCEEWVC-DESK------------DALEELEGFFSKEFGLDASEGEL 190

Qy     182 -RN----------------------------CIAYTSPWSPCSTSCGLGVSTRISNVN 210
            ||                            ||  |: || || :|| |:|||::| |
Db     191 TRNNELIAIVKGGLKMLPVFGSEPQSRAFENPKCIVQTTSWSQCSKTCGTGISTRVTNDN 250

Qy     211 AQCWPEQESRLCNLRPCDVDIHTLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCG 270
             |  :|:|:| :|||    :  :| ||||    :  : : || ||| | : |:|||||
Db     251 PDCKLIKETRICEVRPCGQPSYASLKKGKKCTKTKKSPSPVRFTYAGCSSVKKYRPKYCG 310

Qy     271 VCMDNRCCIPYKSKTIDVSFQCPDGLGFSRQVLWINACFCNLSCRNPNDIFADLESYPDF 330
           |:| ||| | :::|: : |:| ||  |:: |: | :| || :| : |      |:|| |
Db     311 SCVDGRCCTPQQTRTVKIRFRCDDGETFTKSVMMIQSCRCNYNCPHAN------EAYP-F 363

Qy     331 SEIAN 335
             : |
Db     364 YRLVN 368


RESULT   5
S20078
NOV protein - chicken
C;Species: Gallus gallus (chicken)
C;Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jul-2000
C;Accession: S20078
R;Joliot, V.; Martinerie, C.; Dambrine, G.; Plassiart, G.; Brisac, M.; Crochet, J.; Perb
Mol. Cell. Biol. 12, 10-21, 1992
A;Title: Proviral rearrangements and overexpression of a new cellular gene (nov) in myel
A;Reference number: S20078; MUID:92107157
A;Accession: S20078
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-351 <JOL>
A;Cross-references: EMBL:X59284; NID:g63702; PIDN:CAA41975.1; PID:g63703
C;Genetics:
A;Gene: NOV


  Query Match             38.4%;  Score 743.5;  DB 2;  Length 351;
  Best Local Similarity   41.2%;  Pred. No. 5.3e-48;
  Matches  135;  Conservative   49;  Mismatches  117;  Indels   27;  Gaps    7;

Qy       9 DTSSRPQFCKWPC--ECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRG 66
           : |:|   |  ||   ||  ||||  ||  : ||| || :||:| |::|:    ||   |
Db      23 EVSGREAACPRPCGGRCPAEPPRCAPGVPAVLDGCGCCLVCARQRGESCSPLLPCDESGG 82

Qy      67 LYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC- 125
           ||||  |       |:|  : |  || ||: | ||::|||:||| ||| || :|| | |
Db      83 LYCD-RGPEDGGGAGICMVLEGDNCVFDGMIYRNGETFQPSCKYQCTCRDGQIGCLPRCN 141

Qy     126 LRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRD---TGAFDAVG------- 175
           | :  |   || ||:: :|| |||:|||:           |||    | |
Db     142 LGLLLPGPDCPFPRKIEVPGECCEKWVCD----------PRDEVLLGGFAMAAYRQEATL 191

Qy     176 --EVEAWHRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHT 233
             :|     |||  |: || || |||:| |||::| | ||   :::||| :|||: :
Db     192 GIDVSDSSANCIEQTTEWSACSKSCGMGFSTRVTNRNQQCEMVKQTRLCMMRPCENE-EP 250

Qy     234 LIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCP 293
           | ||||:   :   :: |    | | :|:|:|:|||:| | ||| |: :||| | |:||
Db     251 SDKKGKKCIQTKKSMKAVRFEYKNCTSVQTYKPRYCGLCNDGRCCTPHNTKTIQVEFRCP 310

Qy     294 DGLGFSRQVLWINACFCNLSCRNPNDIF 321
           |    : :: || | |: :|   |: |
Db     311 QGKFLKKPMMLINTCVCHGNCPQSNNAF 338


RESULT   6
I38069
gene novH protein - human
C;Species: Homo sapiens (man)
C;Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 05-Nov-1999
C;Accession: I38069
R;Martinerie, C.; Huff, V.; Joubert, I.; Badzioch, M.; Saunders, G.; Strong, L.; Perb
Oncogene 9, 2729-2732, 1994
A;Title: Structural analysis of the human nov proto-oncogene and expression in Wilms
A;Reference number: I38069; MUID:94336229
A;Accession: I38069
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-357 <RES>
A;Cross-references: EMBL:X78351; NID:g587422; PIDN:CAA55146.1; PID:g825696
C;Genetics:
A;Gene: novH
A;Introns: 28/3; 104/1; 188/1; 259/3
C;Superfamily: thrombospondin type 1 repeat homology
F;203-250/Domain: thrombospondin type 1 repeat homology <THR1>


  Query Match             37.5%;  Score 726.5;  DB 2;  Length 357;
  Best Local Similarity   42.0%;  Pred. No. 9.7e-47;
  Matches  137;  Conservative   46;  Mismatches  116;  Indels   27;  Gaps    8;

Qy      10 TSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGLYC 69
           |   |  |  |  || :|| |  ||  : ||| || :||:| |::|::   ||   ||||
Db      32 TQRCPPQC--PGRCPATPPTCAPGVRAVLDGCSCCLVCARQRGESCSDLEPCDESSGLYC 89

Qy      70 DYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-LRV 128
           | | | |    |:|  | |  || ||| | :|: |||:||: ||| || :|| | | | |
Db      90 DRSAD-PSNQTGICTAVEGDNCVFDGVIYRSGEKFQPSCKFQCTCRDGQIGCVPRCQLDV 148

Qy     129 RPPRLWCPHPRRVSIPGHCCEQWVC---EDD--------AKRPRKTAPRDTGAFDAVG-E 176
           |   || ||:| :|| |||:|:|   |:|         | ||  |           :| |
Db     149 LLPEPNCPAPRKVEVPGECCEKWICGPDEEDSLGGLTLAAYRPEAT----------LGVE 198

Qy     177 VEAWHRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVD-IHTLI 235
           |     |||  |: |: || |||:| |||::| | ||   :::||| :|||: :
Db     199 VSDSSVNCIEQTTEWTACSKSCGMGFSTRVTNRNRQCEMLKQTRLCMVRPCEQEPEQPTD 258

Qy     236 KAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDG 295
           | |||||   :   :::     | |  :|:|::|||| | ||| |: :|||   ||| |
Db     259 KKGKKCLRTKKSLKAIHLQFKNCTSLHTYKPRFCGVCSDGRCCTPHNTKTIQAEFQCSPG 318

Qy     296 LGFSRQVLWINACFCNLSCRNPNDIF 321
             : |: |  | |: :|   |: |
Db     319 QIVKKPVMVIGTCTCHTNCPKNNEAF 344


RESULT   7
A43932
mucin 2 precursor, intestinal - human (fragments)
N;Alternate names: mucin SMUC-41
C;Species: Homo sapiens (man)
C;Date: 10-Mar-1993 #sequence_revision 12-Apr-1996 #text_change 05-Nov-1999
C;Accession: A49963; A45106; B45106; A43932; B33532; A61257; PQ0328; PQ0329
R;Gum Jr., J.R.; Hicks, J.W.; Toribara, N.W.; Siddiki, B.; Kim, Y.S.
J. Biol. Chem. 269, 2440-2446, 1994
A;Title: Molecular cloning of human intestinal mucin (MUC2) cDNA. Identification of t
A;Reference number: A49963; MUID:94132002
A;Accession: A49963
A;Molecule type: mRNA
A;Residues: 1-639 <GU1>
A;Cross-references: GB:L21998
R;Gum Jr., J.R.; Hicks, J.W.; Toribara, N.W.; Rothe, E.M.; Lagace, R.E.; Kim, Y.S.
J. Biol. Chem. 267, 21375-21383, 1992
A;Title: The human MUC2 intestinal mucin has cysteine-rich subdomains located both up
A;Reference number: A45106; MUID:93016075
A;Accession: A45106
A;Status: not compared with conceptual translation
```

```
A;Molecule type: mRNA
A;Residues: 626-1895 <GU2>
A;Cross-references: GB:M94131; NID:g186395; PIDN:AAA59163.1; PID:g186396
A;Note: sequence extracted from NCBI backbone (NCBIP:116706)
A;Accession: B45106
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 2037-3020 <GU3>
A;Cross-references: GB:M94132; NID:g186397; PIDN:AAA59164.1; PID:g186398
A;Experimental source: colon
A;Note: sequence extracted from NCBI backbone (NCBIP:116698)
R;Toribara, N.W.; Gum Jr., J.R.; Culhane, P.J.; Lagace, R.E.; Hicks, J.W.; Petersen, G.M
J. Clin. Invest. 88, 1005-1013, 1991
A;Title: MUC-2 human small intestinal mucin gene structure. Repeated arrays and polymorp
A;Reference number: A43932; MUID:91358717
A;Accession: A43932
A;Molecule type: DNA
A;Residues: 1343-1350,'L',1352-1411,'S',1413-1448,'P',1450-1503,'T',1505-1915 <TOR>
A;Cross-references: GB:M74027; NID:g188863; PIDN:AAA59875.1; PID:g188864
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:55749, NCBIP:55750)
R;Gum, J.R.; Byrd, J.C.; Hicks, J.W.; Toribara, N.W.; Lamport, D.T.A.; Kim, Y.S.
J. Biol. Chem. 264, 6480-6487, 1989
A;Title: Molecular cloning of human intestinal mucin cDNAs. Sequence analysis and eviden
A;Reference number: A33532; MUID:89197956
A;Accession: B33532
A;Molecule type: mRNA
A;Residues: 1916-2193 <GU4>
A;Cross-references: GB:M22405; NID:g188873; PIDN:AAA36334.1; PID:g188874
A;Experimental source: intestine
R;Jany, B.H.; Gallup, M.W.; Yan, P.S.; Gum, J.R.; Kim, Y.S.; Basbaum, C.B.
J. Clin. Invest. 87, 77-82, 1991
A;Title: Human bronchus and intestine express the same mucin gene.
A;Reference number: A61257; MUID:91086481
A;Accession: A61257
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 'T',1925-1948,'TTS',1952-1954 <JAN>
A;Experimental source: bronchus
R;Xu, G.; Huan, L.; Khatri, I.; Sajjan, U.S.; McCool, D.; Wang, D.; Jones, C.; Forstner,
Biochem. Biophys. Res. Commun. 183, 821-828, 1992
A;Title: Human intestinal mucin-like protein (MLP) is homologous with rat MLP in the C-t
A;Reference number: PQ0328; MUID:92198477
A;Accession: PQ0328
A;Molecule type: mRNA
A;Residues: 2328-2468 <XUG>
A;Cross-references: GB:M86523
A;Experimental source: small intestine
A;Accession: PQ0329
A;Molecule type: protein
A;Residues: 2328-2342,'K',2344-2354 <XUG1>
C;Genetics:
A;Gene: GDB:MUC2
A;Cross-references: GDB:120203; OMIM:158370
A;Map position: 11p15.5-11p15.5
C;Superfamily: von Willebrand factor; von Willebrand factor type A repeat homology; von
C;Keywords: glycoprotein; intestine; tandem repeat
F;2766-2834/Domain: von Willebrand factor type C repeat homology <VWC>


  Query Match             7.8%;  Score 151.5;  DB 2;  Length 3020;
  Best Local Similarity   20.5%;  Pred. No. 0.002;
  Matches   99;  Conservative   40;  Mismatches  156;  Indels  189;  Gaps   26;

Qy       1 DFTPAPLEDTSSRPQFCKWPCECPPSP--PRCPLGVSLITDGC-------ECCK------ 45
           | ||:||        | :     ||      |      : | |       ||
Db    2527 DCTPSPLCQLIKDSLFAQCHALVPPQHYYDAC------VFDSCFMPGSSLECASLQAYAA 2580

Qy      46 MCAQQL----GDNCTEAAI---CDPHR-------------------------GLYCDY 71
           :||||       | |  |    |  ||                        | :|
Db    2581 LCAQQNICLDWRNHTHGACLVECPSHREYQACGPAEEPTCKSSSSQQNNTVLVEGCFCPE 2640

Qy      72 SGDRPRYAIGVCAQVVGVGCV-LDGVRYNNGQSFQPNCKYNCTCIDGAVG--CTP----- 123
                 || |    |  |||  | |    |: |: :|| || |::|  |  | |
Db    2641 G--TMNYAPGFDVCVKTCGCVGPDNVPREFGEHFEFDCK-NCVCLEGGSGIICQPKRCSQ 2697

Qy     124 ----------LCLRVRP-------------------PRLWCP---HPRRVSIPGHCC 148
                    |  | |                     | : ||     :   :|| ||
Db    2698 KPVTHCVEDGTYLATEVNPADTCCNITVCKCNTSLCKEKPSV-CPLGFEVKSKMVPGRCC 2756

Qy     149 EQWVCEDDAKRPRKTAPRDTGA-------FDAVGEVEAWHR---NCIAYTSPWSPCSTSC 198
             : ||         |    |:       | |   :  :     | || |     ||:|||
Db    2757 PFYWCESKGVCVHGNAEYQPGSPVYSSKCQDCVCTDKVDNNTLLNVIACTH--VPCNTSC 2814

Qy     199 GLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGK-----------KCLAVYQ 246
             |    :     :|  : |   | ::  | : | ::| |                |: ::
Db    2815 SPGF--ELMEAPGECCKKCEQTHCIIKRPD-NQHVILKPGDFKSDPKNNCTFFSCVKIHN 2871

Qy     247 PEASM-------NFTLAGCI-STRSYQPKYC---------------------------- 269
             |         ||  : ||  : :: |  |
Db    2872 QLISSVSNITCPNFDASICIPGSITFMPNGCCKTCTPRNETRVPCSTVPVTTEVSYAGCT 2931

Qy     270 ---------GVC------------MDN--RCCIPYKSKTIDVSFQCPDGLGFSRQVLWIN 306
                    | |            :|:   ||   |:   :|   ||:|   :     |
Db    2932 KTVLMNHCSGSCGTFVMYSAKAQALDHSCSCCKEEKTSQREVVLSCPNGGSLTHTYTHIE 2991

Qy     307 ACFC 310
           :| |
Db    2992 SCQC 2995


RESULT   8
A46019
Notch-1 protein - mouse
N;Alternate names: motch protein
C;Species: Mus musculus (house mouse)
C;Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
C;Accession: A46019; S25144
R;del Amo, F.F.; Gendron-Maguire, M.; Swiatek, P.J.; Jenkins, N.A.; Copeland, N.G.; G
Genomics 15, 259-264, 1993
A;Title: Cloning, analysis, and chromosomal localization of Notch-1, a mouse homolog
A;Reference number: A46019; MUID:93194170
A;Accession: A46019
A;Status: not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-2531 <DEL>
A;Cross-references: GB:Z11886; GB:S47228; NID:g288502; PIDN:CAA77941.1; PID:g288503
A;Note: sequence extracted from NCBI backbone (NCBIP:127318)
R;Franco del Amo, F.; Smith, D.E.; Swiatek, P.J.; Gendron-Maguire, M.; Greenspan, R.J
submitted to the EMBL Data Library, April 1992
A;Description: Expression pattern of Motch, a mouse homolog of Drosophila Notch, sugg
A;Reference number: S25144
A;Accession: S25144
A;Molecule type: mRNA
A;Residues: 1551-2108,'Q',2110-2114,'ALP',2118-2170 <FRA>
A;Cross-references: EMBL:Z11886
C;Genetics:
A;Gene: notch-1
A;Map position: 2
A;Note: proximal region of chromosome 2
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol
F;106-138/Domain: EGF homology <EGF1>
F;144-175/Domain: EGF homology <EG01>
F;222-254/Domain: EGF homology <EGF2>
F;261-292/Domain: EGF homology <EG02>
F;339-370/Domain: EGF homology <EG03>
F;416-449/Domain: EGF homology <EGF3>
F;456-487/Domain: EGF homology <EG04>
F;494-525/Domain: EGF homology <EG05>
F;532-563/Domain: EGF homology <EG06>
F;607-638/Domain: EGF homology <EG07>
F;682-713/Domain: EGF homology <EG08>
F;757-788/Domain: EGF homology <EG09>
F;795-826/Domain: EGF homology <EG10>
F;873-904/Domain: EGF homology <EG11>
F;911-942/Domain: EGF homology <EG12>
```

F;949-980/Domain: EGF homology <EG13>
F;987-1018/Domain: EGF homology <EG14>
F;1025-1056/Domain: EGF homology <EG15>
F;1063-1094/Domain: EGF homology <EG16>
F;1149-1180/Domain: EGF homology <EG17>
F;1187-1218/Domain: EGF homology <EG18>
F;1233-1264/Domain: EGF homology <EGF4>
F;1352-1383/Domain: EGF homology <EG19>
F;1391-1425/Domain: EGF homology <EGF>
F;1917-1948/Domain: ankyrin repeat homology <AN1>
F;1949-1981/Domain: ankyrin repeat homology <AN2>
F;1983-2015/Domain: ankyrin repeat homology <AN3>
F;2016-2048/Domain: ankyrin repeat homology <AN4>
F;2049-2081/Domain: ankyrin repeat homology <AN5>

Query Match 7.7%; Score 150; DB 2; Length 2531;
Best Local Similarity 21.6%; Pred. No. 0.0023;
Matches 90; Conservative 31; Mismatches 135; Indels 160; Gaps 26;

```
Qy      21 CECPPS--PPRCPLGVS-------LITDGCEC--------CKMCAQQLGDNCTEAAI-CD 62
           | |||    | |   |:       |   |  |         |  ||   | :|    : |
Db     164 CRCPPGFHGPTCRQDVNECSQNPGLCRHGGHCHNEIGSYRCACCATHTGPHCELPYVPCS 223

Qy      63 P---HRGLYCDYSGDR-------PRYAIGVCAQVV----------GVGCVLDGVR-YN-- 99
           |     |  |  :||         | :|   | : |          |  || |||  ||
Db     224 PSPCQNGATCRPTGDTTHECACLPGFAGQNCEENVDDCPGNNCKNGGACV-DGVNTYNCR 282

Qy     100 -----NGQ---------SFQPN-CK-----------YNCTCIDGAVG----------CTP 123
                 ||            || |:           ||| |::|  |           :
Db     283 CPPEVTGQYCTEDVDECQLMPNACQNAGTCHNTHGGYNCVCVNGWTGEDCSENIDDCASA 342

Qy     124 LCL-------RVRPPRLWCPHPR--------RVSIPGHCCEQWVCEDDAKRPRKTAPRDT 168
           |        ||      ||| |          |   | |   |: :    ::     :
Db     343 ACFQGATCHDRVASFYCECPHGRTGLLCHLKHACISNPCNEGSNCDTNPVNGKRICTCPS 402

Qy     169 GAFDAVGEVEAWHRNCIAYTSPWSPCS---TSCGLGVS---------TRISNVNAQCWP- 215
           |                 || |   ||     | || :          : :   ||
Db     403 G-----------------YTGP--ACSQDVDECDLGANRCEHAGKCLNTLGSFECQCLQG 443

Qy     216 ------EQESRLCNLRPCDVDIHTLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYC 269
                | :   |   ||  |   | : |:     :|           ||    |:  ||
Db     444 YTGPGCEIDVNECISNPCQNDATCLDQIGE-----FQ----------CICMPGYEGVYC 487

Qy     270 GV---------CMDNRCCIPYKSKTIDVSFQCPDGLGFSRQVLWINACFCNLSCRN 316
           :         |: |  |:    |  :   ||| |         :: |  :  |:|
Db     488 EINTDECASSPCLHNGHCM---DKIHEFQCQCPKGFNGHLCQYDVDEC-ASTPCKN 539
```

RESULT 9
VWHU
von Willebrand factor precursor - human
C;Species: Homo sapiens (man)
C;Date: 04-Dec-1986 #sequence_revision 30-Jun-1993 #text_change 22-Jun-1999
C;Accession: A34480; S02377; A37139; S23676; A25298; A25469; A25366; S23618; S23645; A94
R;Mancuso, D.J.; Tuley, E.A.; Westfield, L.A.; Worrall, N.K.; Shelton-Inloes, B.B.; Sora
J. Biol. Chem. 264, 19514-19527, 1989
A;Title: Structure of the gene for human von Willebrand factor.
A;Reference number: A34480; MUID:90062044
A;Accession: A34480
A;Molecule type: DNA
A;Residues: 1-2813 <MAN>
A;Cross-references: EMBL:M25864
R;Bonthron, D.; Orkin, S.H.
Eur. J. Biochem. 171, 51-57, 1988
A;Title: The human von Willebrand factor gene. Structure of the 5' region.
A;Reference number: S02377; MUID:88111704
A;Accession: S02377
A;Molecule type: DNA
A;Residues: 1-177 <BO2>
A;Cross-references: EMBL:X06828
R;Mancuso, D.J.; Tuley, E.A.; Westfield, L.A.; Lester-Mancuso, T.L.; Le Beau, M.M.; Sora
Biochemistry 30, 253-269, 1991
A;Title: Human von Willebrand factor gene and pseudogene: structural analysis and dif
A;Reference number: A37139; MUID:91105089
A;Accession: A37139
A;Molecule type: DNA
A;Residues: 990-1947 <MAD>
A;Cross-references: GB:M60675; NID:g340357; PIDN:AAA61295.1; PID:g553810
A;Note: the authors translated the codon CGC for residue 156 as Gln
R;Collins, C.J.; Underdahl, J.P.; Levene, R.B.; Ravera, C.P.; Morin, M.J.; Dombalagia
Proc. Natl. Acad. Sci. U.S.A. 84, 4393-4397, 1987
A;Title: Molecular cloning of the human gene for von Willebrand factor and identifica
A;Reference number: S23676; MUID:87260814
A;Accession: S23676
A;Molecule type: DNA
A;Residues: 2731-2813 <COL>
A;Cross-references: EMBL:M16945
R;Bonthron, D.; Orr, E.C.; Mitsock, L.M.; Ginsburg, D.; Handin, R.I.; Orkin, S.H.
Nucleic Acids Res. 14, 7125-7127, 1986
A;Title: Nucleotide sequence of pre-pro-von Willebrand factor cDNA.
A;Reference number: A25298; MUID:87016349
A;Accession: A25298
A;Molecule type: mRNA
A;Residues: 1-470,'V',472-2813 <BON>
A;Cross-references: EMBL:X04385
R;Verweij, C.L.; Diergaarde, P.J.; Hart, M.; Pannekoek, H.
EMBO J. 5, 1839-1847, 1986
A;Title: Full-length von Willebrand factor (vWF) cDNA encodes a highly repetitive pro
A;Reference number: A91044; MUID:87004550
A;Accession: A25469
A;Molecule type: mRNA
A;Residues: 1-470,'V',472-483,'R',485-1022,'K',1024-1025,'E',1027-1400 <VER>
A;Cross-references: EMBL:X04146
A;Note: this sequence has been revised in reference A91056
R;Verweij, C.L.; Diergaarde, P.J.; Hart, M.; Pannekoek, H.
EMBO J. 5, 3074, 1986
A;Reference number: A91056
A;Accession: A25366
A;Molecule type: mRNA
A;Residues: 1021-1030 <VE2>
A;Note: this is a revision to the sequence from reference A91044
R;Shelton-Inloes, B.B.; Broze Jr., G.J.; Miletich, J.P.; Sadler, J.E.
Biochem. Biophys. Res. Commun. 144, 657-665, 1987
A;Title: Evolution of human von Willebrand factor: cDNA sequence polymorphisms, repea
A;Reference number: S23618; MUID:87213253
A;Accession: S23618
A;Molecule type: mRNA
A;Residues: 1-120 <SH2>
A;Cross-references: EMBL:M17588; NID:g799330; PIDN:AAA65940.1; PID:g340316
A;Accession: S23645
A;Molecule type: protein
A;Residues: 23-56 <SH3>
R;Sadler, J.E.; Shelton-Inloes, B.B.; Sorace, J.M.; Harlan, J.M.; Titani, K.; Davie,
Proc. Natl. Acad. Sci. U.S.A. 82, 6394-6398, 1985
A;Title: Cloning and characterization of two cDNAs coding for human von Willebrand fa
A;Reference number: A94060; MUID:86016708
A;Accession: A94060
A;Molecule type: mRNA
A;Residues: 'WA',739,'C',744-769,'H',771-788,'A',790-803,'S',805-873;1289-1471,'D',14
A;Note: the authors translated the codon TCG for residue 2168 as Cys
R;Shelton-Inloes, B.B.; Titani, K.; Sadler, J.E.
Biochemistry 25, 3164-3171, 1986
A;Title: cDNA sequences for human von Willebrand factor reveal five types of repeated
A;Reference number: A90504; MUID:86269894
A;Accession: A90504
A;Molecule type: mRNA
A;Residues: 781-788,'A',790-1424 <SHE>
A;Note: 852-Gln, 857-Asp, and 1381-Thr were also found
R;Ginsburg, D.; Handin, R.I.; Bonthron, D.T.; Donlon, T.A.; Bruns, G.A.P.; Latt, S.A.
Science 228, 1401-1406, 1985
A;Title: Human von Willebrand factor (vWF): isolation of complementary DNA (cDNA) clo
A;Reference number: A44178; MUID:85244588
A;Accession: A44178
A;Molecule type: mRNA
A;Residues: 2621-2813 <GIN>

```
A;Cross-references: EMBL:K03028; NID:g340308; PIDN:AAA61293.1; PID:g340309
R;Verweij, C.L.; de Vries, C.J.M.; Distel, B.; van Zonneveld, A.J.; van Kessel, A.G.; va
Nucleic Acids Res. 13, 4699-4717, 1985
A;Title: Construction of cDNA coding for human von Willebrand factor using antibody prob
A;Reference number: S07363; MUID:85269603
A;Accession: S07363
A;Molecule type: mRNA
A;Residues: 2731-2813 <VE3>
A;Cross-references: EMBL:X02672; NID:g37939; PIDN:CAA26503.1; PID:g37940
R;Lynch, D.C.; Zimmerman, T.S.; Collins, C.J.; Brown, M.; Morin, M.J.; Ling, E.H.; Livin
Cell 41, 49-56, 1985
A;Title: Molecular cloning of cDNA for human von Willebrand factor: authentication by a
A;Reference number: S23678; MUID:85201687
A;Accession: S23678
A;Molecule type: mRNA
A;Residues: 2731-2813 <LYN>
A;Cross-references: EMBL:K03028
R;Titani, K.; Kumar, S.; Takio, K.; Ericsson, L.H.; Wade, R.D.; Ashida, K.; Walsh, K.A.;
Biochemistry 25, 3171-3184, 1986
A;Title: Amino acid sequences of human von Willebrand factor.
A;Reference number: A90505; MUID:86269895
A;Accession: A90505
A;Molecule type: protein
A;Residues: 764-788,'A',790-1471,'D',1473-2813 <TIT>
A;Note: 789-Thr was also found
R;Chopek, M.W.; Girma, J.P.; Fujikawa, K.; Davie, E.W.; Titani, K.
Biochemistry 25, 3146-3155, 1986
A;Title: Human von Willebrand factor: a multivalent protein composed of identical subuni
A;Reference number: A23464; MUID:86269892
A;Accession: A23464
A;Molecule type: protein
A;Residues: 764-773;2803-2813 <CHO>
R;Dent, J.A.; Berkowitz, S.D.; Ware, J.; Kasper, C.K.; Ruggeri, Z.M.
Proc. Natl. Acad. Sci. U.S.A. 87, 6306-6310, 1990
A;Title: Identification of a cleavage site directing the immunochemical detection of mol
A;Reference number: A36013; MUID:90349604
A;Accession: A36013
A;Molecule type: protein
A;Residues: 1606-1617 <DEN>
R;Fay, P.J.; Kawai, Y.; Wagner, D.D.; Ginsburg, D.; Bonthron, D.; Ohlsson-Wilhelm, B.M.;
Science 232, 995-998, 1986
A;Title: Propolypeptide of von Willebrand factor circulates in blood and is identical to
A;Reference number: A60913; MUID:86208144
A;Accession: A60913
A;Molecule type: protein
A;Residues: 576-590 <FAY>
C;Genetics:
A;Gene: GDB:VWF
A;Cross-references: GDB:119125; OMIM:193400
A;Map position: 12p13.3-12p13.2
A;Introns: 19/1; 74/1; 108/2; 178/1; 219/3; 292/1; 333/1; 370/2; 386/1; 431/3; 478/1; 51
5/1; 1724/1; 1771/1; 1819/1; 1874/1; 1888/3; 1948/1; 2021/3; 2086/1; 2200/1; 2266/3; 230
C;Superfamily: von Willebrand factor; von Willebrand factor type A repeat homology; von
C;Keywords: blood coagulation; cell binding; connective tissue; disulfide bond; duplicat
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-763/Product: von Willebrand antigen II #status predicted <MA1>
F;34-386/Domain: type D repeat 1 <DD1>
F;387-745/Domain: type D repeat 2 <DD2>
F;698-700/Region: cell attachment (R-G-D) motif
F;764-2813/Product: von Willebrand factor #status predicted <MA2>
F;784-865/Domain: D' <DDD>
F;788-833,2216-2261/Region: duplication
F;826-853,2400-2515,2544-2662/Region: duplication
F;842-1130,1934-2203/Region: duplication
F;866-1241/Domain: type D repeat 3 <DD3>
F;1275-1443/Domain: von Willebrand factor type A repeat homology <VWA1>
F;1496-1654/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1689-1854/Domain: von Willebrand factor type A repeat homology <VWA3>
F;1947-2295/Domain: type D repeat 4 <DD4>
F;2296-2330/Domain: type B repeat 1 <VB1>
F;2340-2365/Domain: type B repeat 2 <VB2>
F;2375-2399/Domain: type B repeat 3 <VB3>
F;2430-2497/Domain: von Willebrand factor type C repeat homology <VWC1>
F;2507-2509/Region: cell attachment (R-G-D) motif
F;2581-2647/Domain: von Willebrand factor type C repeat homology <VWC2>
F;857,1231,1515,1574,2223,2290,2357,2400,2546,2585,2790/Binding site: carbohydrate (A
F;1147/Binding site: carbohydrate (Asn) (covalent) #status atypical
F;1248,1255,1256,1468,1477,1487,1679,2298/Binding site: carbohydrate (Thr) (covalent)
F;1263,1486/Binding site: carbohydrate (Ser) (covalent) #status experimental


  Query Match             7.7%;  Score 150;  DB 1;  Length 2813;
  Best Local Similarity   22.9%;  Pred. No. 0.0025;
  Matches   93;  Conservative   43;  Mismatches  112;  Indels  158;  Gaps   31;

Qy      14 PQFCKWPCECPPSPPRCPLGVSLITDGCE--CCKMC---AQQLGDNCTEAAICDPHRGLY 68
           | ||     |||          ||: : ||  | : |       ||: :|   | | : :
Db    2196 PDFC--AMSCPP---------SLVYNHCEHGCPRHCDGNVSSCGDHPSEGCFCPPDKVML 2244

Qy      69 ---CDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPN---CKYNCTCIDG-AVGC 121
             |      | |   | | :|     |||::   ::: |:   |:  |||: |  | |
Db    2245 EGSC-----VPEEA---CTQCIGE----DGVQHQFLEAWVPDHQPCQI-CTCLSGRKVNC 2291

Qy     122 T---------PLC-----LRVRP------PRLWC-----------PHPRR-----VSIP 144
           |         | |      |:|       |   |           ||  |     :: |
Db    2292 TTQPCPTAKAPTCGLCEVARLRQNADQCCPEYECVCDPVSCDLPPVPHCERGLQPTLTNP 2351

Qy     145 GHCCEQWVC---EDDAKR--PRKTAP------RDTGAFDAVGEVEAWHRNCIAYTSPWSP 193
           | |   : |   ::: ||  |     |      | |   |   | |    ||:  |
Db    2352 GECRPNFTCACRKEECKRVSPPSCPPHRLPTLRKTQCCD---EYEC-ACNCVNST----- 2402

Qy     194 CSTSCGLG--VSTRISN---VNAQCWPEQESRLCNLRP-----------CDVDIHTLIK 236
             || ||   ||  ::       | |:   ::|  |               |||
Db    2403 --VSCPLGYLASTATNDCGCTTTTCLPD---KVCVHRSTIYPVGQFWEEGCDV------- 2450

Qy     237 AGKKCLAVYQPEASMNFTLAGCI------STRS------YQPKYCGVCMDNRCCIPYKS- 283
             |      :| |   :| |       | ||      :: : || |: : | :   |
Db    2451 ----CTCTDMEDAVMGLRVAQCSQKPCEDSCRSGFTYVLHEGECCGRCLPSACEVVTGSP 2506

Qy     284 --------KTIDVSFQCPDGLGFSRQVLWINACFCNLSCRNPNDIF 321
                  |::   :  |:           | | |   |   ::|
Db    2507 RGDSQSSWKSVGSQWASPE----------NPCLINECVRVKEEVF 2541


RESULT  10
A39804
thrombospondin precursor - chicken
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A39804
R;Lawler, J.; Duquette, M.; Ferro, P.
J. Biol. Chem. 266, 8039-8043, 1991
A;Title: Cloning and sequencing of chicken thrombospondin.
A;Reference number: A39804; MUID:91217026
A;Accession: A39804
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1178 <LAW>
A;Cross-references: GB:M60853; NID:g212763; PIDN:AAA51437.1; PID:g212764
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology;
F;325-383/Domain: von Willebrand factor type C repeat homology <VWC>
F;386-437/Domain: thrombospondin type 1 repeat homology <THR1>
F;442-498/Domain: thrombospondin type 1 repeat homology <THR2>
F;499-555/Domain: thrombospondin type 1 repeat homology <THR3>
F;658-697/Domain: EGF homology <EGF>


  Query Match             7.6%;  Score 147.5;  DB 1;  Length 1178;
  Best Local Similarity   21.8%;  Pred. No. 0.0018;
  Matches   86;  Conservative   38;  Mismatches  121;  Indels  149;  Gaps   23;

Qy       8 EDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAA-------- 59
           | |  : :||   ||          || :: |:     ::    | ||  : :
Db     263 EKTEKKAEFCDRSCE--------ELG-TMFTE-LTGLRIVVNNLADNLQKVSEENQIMWE 312

Qy      60 ICDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAV 119
```

```
          :  |:: |             || |        ||  : : :|:  :    ||| |  :
Db    313 LIGPNKTLKNQ----------SVCWQ--------DGRVFADSESWIVDSCTKCTCQDSKI 354

Qy    120 GCTPLCLRVRPPRLWCPHPRRVSIPGHCC----------EQW-----------VCED--- 155
          :| ::  | : |  |    | | ||          | |           |
Db    355 ----VCHQITCPPVSCADPS--FIEGECCPVCSHSDDSEEGWSPWSDWTKCSVTCGSGTQ 408

Qy    156 ------DAKRPRKTAP---------------RDTGAFDAVGEVEAWHRNCIAYTSPWSP 193
               |  |   | |                |  |          |     :: ||||
Db    409 MRGRSCDVTRSACTGPHIQTRMCSFKKCDHRIRQDG---------GW-----SHWSPWSS 454

Qy    194 CSTSCGLGVSTRISNVNAQCWPE----------QESRLCNLRPCDVDIHTLIKAGKKCLA 243
          || :||:|  |||   |:   |:          :|:  |   || |:       |:
Db    455 CSVTCGVGNITRIRLCNSPI-PQMGGKNCVGNGRETEKCEKAPCPVN-------GQ--WG 504

Qy    244 VYQPEASMNFTLAGCISTRSY-----QPKYCGVCMDNRCCIPYKSKTIDVSFQCP-DGLG 297
           : | ::   |  | | ||      :|:| |      |   |   :     || ||
Db    505 PWSPWSACTVTCGGGIRERSRLCNSPEPQYGG----KPCVGDTKQHDMCNKRDCPIDG-- 558

Qy    298 FSRQVLWINACFCNLSCRNPNDIFADLESYPDFS 331
               | ||    |           |||| |
Db    559 -----CLSNPCFPGAEC----------NSYPDGS 577


RESULT  11
A57534
mucin 5AC (clone L31) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 20-Apr-2000
C;Accession: A57534
R;Lesuffleur, T.; Roche, F.; Hill, A.S.; Lacasa, M.; Fox, M.; Swallow, D.M.; Zweibaum, A
J. Biol. Chem. 270, 13665-13673, 1995
A;Title: Characterization of a mucin cDNA clone isolated from HT-29 mucus-secreting cell
A;Reference number: A57534; MUID:95293957
A;Accession: A57534
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1042 <LES>
A;Cross-references: GB:Z48314; NID:g1052607; PIDN:CAA88307.1; PID:g1052608
C;Genetics:
A;Gene: GDB:MUC5AC
A;Cross-references: GDB:454136; OMIM:158373
A;Map position: 11p15.5-11p15.5
C;Superfamily: von Willebrand factor type C repeat homology
F;678-746/Domain: von Willebrand factor type C repeat homology <VWC>


  Query Match             7.5%;  Score 145.5;  DB 2;  Length 1042;
  Best Local Similarity   20.9%;  Pred. No. 0.0023;
  Matches   84;  Conservative   29;  Mismatches  137;  Indels  151;  Gaps   23;

Qy     13 RPQFCKWPCECPPSPPRCPL----GVSLITDGCECC---------KMCAQQLGDNCTEAA 59
          ||: |       | || |||     |      :||           |   :|  | | |
Db    607 RPKLC-------PLPPACPLPGFVPVPAAPQAGQCCPQYSCACNTSRCPAPVG--CPEGA 657

Qy     60 ICDPHRGLYCDYSGDRPRYAIGVCAQVVGVG---CVLDGVRYNNGQSFQPNCKYNCTC-I 115
          | :       | |  | |  |       | ::|  |  |     :    | | :
Db    658 -----RAI--------PTYQEGACCPVQNCSWTVCSINGTLYQPGAVVSSSLCETCRCEL 704

Qy    116 DG-------AVGC-TPLCLRVRPPRLWCPHPRRVSI-----PGHCCEQWVCEDDAKRPRK 162
           |       | | | :|         ||    |       | ||    |   |
Db    705 PGGPPSDAFVVSCETQIC------NTHCP----VGFEYQEQSGQCC--GTCVQVA----- 747

Qy    163 TAPRDTGAFDAVGEVEAWHRNCIAYTS-----------PWSPCSTSCGLGVSTRISNVNA 211
                         |: ||          ||     |           |
Db    748 --------------------CVTNTSKSPAHLFYPGETWSDAGNHC----------VTH 776

Qy    212 QCWPEQE-------SRLCNLRPCDVDIHTLIKAG-------------KKCLAVYQPEAS 250
          ||   |:       : |    | :|   : | |             :   |||    |
Db    777 QCEKHQDGLVVVTTKKACPPLSCSLDEARMSKDGCCRFCPLPPPPYQNQSTCAVY--HRS 834

Qy    251 MNFTLAGCISTRSYQPKYC-GVCMDN--------------RCCIPYKSKTIDVSFQCPD 294
          :    || |:   :  || | | |:              :||   ::   :|:  | |
Db    835 LIIQQQGCSSSEPVRLAYCRGNCGDSSSMYSLEGNTVEHRCQCCQELRTSLRNVTLHCTD 894

Qy    295 GLGFSRQVLWINACFC-NLSCRNPNDIFADLESYPDFSEIA 334
          |   :     :  | |    |  | |     |: |: |: | |
Db    895 GSSRAFSYTEVEECGCMGRRCPAPGDTQHSEEAEPEPSQEA 935


RESULT  12
A42112
mucin-like peptide MLP 2677 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 10-Oct-1997
C;Accession: A42112
R;Xu, G.; Huan, L.J.; Khatri, I.A.; Wang, D.; Bennick, A.; Fahim, R.E.; Forstner, G.G
J. Biol. Chem. 267, 5401-5407, 1992
A;Title: cDNA for the carboxyl-terminal region of a rat intestinal mucin-like peptide
A;Reference number: A42112; MUID:92184794
A;Accession: A42112
A;Status: preliminary
A;Molecule type: nucleic acid; protein
A;Residues: 1-837 <XU1>
A;Experimental source: intestine
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIP:87420)
C;Superfamily: von Willebrand factor type C repeat homology
F;582-650/Domain: von Willebrand factor type C repeat homology <VWC>


  Query Match             7.5%;  Score 144.5;  DB 2;  Length 837;
  Best Local Similarity   20.1%;  Pred. No. 0.0022;
  Matches   84;  Conservative   37;  Mismatches  133;  Indels  163;  Gaps   25;

Qy     23 CPPS-PPRC----PLGVSLITDGC-------------ECCKMCA--------QQLGDN- 54
          | ||  | |    |   :|: :||              | |:|         :: |::
Db    429 CGPSEEPTCQSSSPKNSTLLVEGCFCPEGTTKFAPGYDVCVKICGCVGPDNVPREFGEHF 488

Qy     55 --------CTEAA---ICDPHR------------GLY-----------------CD--- 70
                 | |      :| | :             | |                 ||
Db    489 EFDCKDCVCLEGGSGIVCQPKKCARGNLTTCEEDGTYLVVEADPDDKCCNTTSCKCDPKR 548

Qy     71 YSGDRP----------RYAIGVCAQVV-----GVGCVLDGVRYNNGQSFQPNCKYNCTCI 115
             :||           :  | |  |      || || :   |  |     |   :| |
Db    549 CKAERPSCLLGFEVKSEHVPGKCCPVYSCVPKGV-CVHENAEYQPGSPVYSNKCQDCVCT 607

Qy    116 DGAVGCTPL----CLRVRPPRLWCPHP-RRVSIPGHCCEQWVCEDD---AKRPR------ 161
          |     | |    |  | | |  : |      | :|| ||::  |:      |||
Db    608 DSMDNSTQLNVISCTHV-PCNISCSSGFELVEVPGECCKK--CQQTHCIIKRPEQQYIIL 664

Qy    162 -----KTAPRDTGAF-----------DAVGEV---EAWHRNCIAYTSPWSP--CSTSCGL 200
               :  | |   |              :|  :   :    :|:  :  : |  |  :|
Db    665 KPGEIQKNPNDRCTFFSCMKINNQLISSVSNITCPDFDPSDCVPGSITYMPNGCCKTC-- 722

Qy    201 GVSTRISNVNAQCWPEQESRLCNLRPC-------DVDIHTLIKAGKKCLAVYQPEASMNF 253
          |  | |             |  ||       ::  :   |             ||||
Db    723 -----IHNPN-----------NTVPCSAIPVMKEISYNGCAK----------NISMNF 754

Qy    254 TLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSRQVLWINACFC 310
          | :   |  :   :     ||   ::    ||  ||||   |     | :| |
Db    755 CAGSCGTFAMYSAQAQDLDHGCSCCREERTSVRMVSLDCPDGSKLSHSYTHIESCLC 811


RESULT  13
T18856
angiogenesis inhibitor homolog - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T18856; T24653
R;McMurray, A.
submitted to the EMBL Data Library, July 1995
A;Reference number: Z19031
A;Accession: T18856
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
```

```
A;Residues: 1-1444 <WIL>
A;Cross-references: EMBL:Z50004; PIDN:CAA90293.1; GSPDB:GN00028; CESP:C02B4.1
A;Experimental source: clone C02B4
R;McMurray, A.
submitted to the EMBL Data Library, July 1995
A;Reference number: Z19917
A;Accession: T24653
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1444 <WI2>
A;Cross-references: EMBL:Z50006; PIDN:CAA90302.1; GSPDB:GN00028; CESP:C02B4.1
A;Experimental source: clone T07C5
C;Genetics:
A;Gene: CESP:C02B4.1
A;Map position: X
A;Introns: 25/3; 70/3; 96/3; 139/3; 187/1; 234/2; 282/3; 376/2; 422/2; 478/3; 509/3; 566

  Query Match             7.4%;  Score 144;  DB 2;  Length 1444;
  Best Local Similarity   25.4%;  Pred. No. 0.0039;
  Matches   70;  Conservative   20;  Mismatches   66;  Indels  120;  Gaps   18;

Qy      10 TSSRPQFC----------------KWPC---ECPPSPPRCPLGVSLITDGCECCKMCAQ 49
           ||:| :||                  : ||    | ||      | || ::   | | |
Db    1149 TSTRRRFCQVVDPTVQGFCAGAILEQIPCAPGSCSPSAG----GWSLWSEWSSCSKDC-- 1202

Qy      50 QLGDN--------CTEAAICDPHRGLYCD-YSGD-RPRYAIGVCAQVVGVGCVLDGVRYN 99
             ||         |:|      :|| ||  || | ||              ||:| |
Db    1203 --GDTGHQIRNRMCSEP--IPSNRGAYCSGYSFDQRP-------------CVMDNV--- 1241

Qy     100 NGQSFQPNCKYNCTCIDGAVG----------CTPLCLRVRPPRL-WCPHPRRVSIPGHC 147
                 | |  |              ||  |      |  :| :|:        |
Db    1242 -------------CSDEKVDGGWTDWTAWSECTDYCRNGHRSRTRFCANPKPSQGGAQC 1287

Qy     148 C----EQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHRNCIAYTSPWSPCSTSCGLGVS 203
           |    | |    | :   || |          |     :  | |:||| ||| ||
Db    1288 TGSDFELNPCFD----PARCHLRDGG----------W-----STWSDWTPCSASCGFGVQ 1328

Qy     204 TRISNVNAQCWPE----------QESRLCNLRPCD 228
           ||  : ::   ||           :: ||:|  ||
Db    1329 TRDRSCSS---PEPKGGQSCSGLAHQTSLCDLPACD 1361


RESULT  14
TSHUP1
thrombospondin 1 precursor - human
C;Species: Homo sapiens (man)
C;Date: 23-Aug-1987 #sequence_revision 03-Aug-1995 #text_change 17-Nov-2000
C;Accession: A26155; A34274; A30140; A25812; A05172; A42927
R;Lawler, J.; Hynes, R.O.
J. Cell Biol. 103, 1635-1648, 1986
A;Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple d
A;Reference number: A26155; MUID:87057617
A;Accession: A26155
A;Molecule type: mRNA
A;Residues: 1-1170 <LAW>
A;Cross-references: GB:X04665; NID:g37137; PIDN:CAA28370.1; PID:g37138
A;Note: parts of this sequence, including the amino end of the mature protein, were dete
R;Laherty, C.D.; Gierman, T.M.; Dixit, V.M.
J. Biol. Chem. 264, 11222-11227, 1989
A;Title: Characterization of the promoter region of the human thrombospondin gene. DNA s
A;Reference number: A34274; MUID:89291870
A;Accession: A34274
A;Molecule type: DNA
A;Residues: 1-166 <LAH>
A;Cross-references: GB:J04835
R;Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein,
J. Cell Biol. 108, 729-736, 1989
A;Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in th
A;Reference number: A30140; MUID:89139590
A;Accession: A30140
A;Molecule type: mRNA
A;Residues: 1-83,'A',85-522,'A',524-1170 <HEN>
A;Cross-references: EMBL:X14787; NID:g37464; PIDN:CAA32889.1; PID:g37465
A;Note: parts of this sequence, including the amino end of the mature protein, were d
R;Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.
Biochemistry 25, 8418-8425, 1986
A;Title: Partial amino acid sequence of human thrombospondin as determined by analysi
A;Reference number: A25812; MUID:87157592
A;Accession: A25812
A;Molecule type: mRNA
A;Residues: 1-83,'A',85-397 <KOB>
A;Cross-references: GB:M25631; NID:g538353; PIDN:AAA36741.1; PID:g538354
R;Dixit, V.M.; Hennessy, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986
A;Reference number: A05172; MUID:86287276
A;Accession: A05172
A;Molecule type: mRNA
A;Residues: 1-83,'A',85-374,'RC' <DIX>
A;Cross-references: GB:M14326; NID:g340005; PIDN:AAA61237.1; PID:g553801
A;Note: parts of this sequence, including the amino end of the mature protein, were d
R;Sun, X.; Skorstengaard, K.; Mosher, D.F.
J. Cell Biol. 118, 693-701, 1992
A;Title: Disulfides modulate RGD-inhibitable cell adhesive activity of thrombospondin
A;Reference number: A42927; MUID:92348511
A;Accession: A42927
A;Molecule type: protein
A;Residues: 987-1003 <SUN>
A;Note: Cys-992 is shown to have a free sulfhydryl
C;Genetics:
A;Gene: GDB:THBS1; TSP1; TSP
A;Cross-references: GDB:120438; OMIM:188060
A;Map position: 15q15-15q15
A;Introns: 23/1
A;Note: the list of introns may be incomplete
C;Complex: homotrimer, disulfide linked
C;Function:
A;Description: participates in cell migration and adhesion, and in platelet aggregati
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology;
C;Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; tri
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-1170/Product: thrombospondin 1 #status predicted <MAT>
F;317-375/Domain: von Willebrand factor type C repeat homology <VWC>
F;378-429/Domain: thrombospondin type 1 repeat homology <THR1>
F;434-490/Domain: thrombospondin type 1 repeat homology <THR2>
F;491-547/Domain: thrombospondin type 1 repeat homology <THR3>
F;551-586/Domain: EGF homology <EGF1>
F;650-689/Domain: EGF homology <EGF2>
F;926-928/Region: cell attachment (R-G-D) motif
F;171-232/Disulfide bonds: #status predicted
F;248,360,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;270,274/Disulfide bonds: interchain #status predicted
F;610/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F;1051/Binding site: carbohydrate (Asn) (covalent) #status absent

  Query Match             7.4%;  Score 142.5;  DB 1;  Length 1170;
  Best Local Similarity   22.6%;  Pred. No. 0.0042;
  Matches   71;  Conservative   30;  Mismatches   78;  Indels  135;  Gaps   20;

Qy      91 CVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLCLRVRPPRLWCPHPRRVSIP-GHCC- 148
           |  :||:| | : :   :    | | :        :| :|  | : | :    ::| | ||
Db     318 CYHNGVQYRNNEEWTVDSCTECHCQNSVT----ICKKVSCPIMPCSN---ATVPDGECCP 370

Qy     149 --------------EQWV----------------------CEDDAKRPR--------K 162
                        :|                        ||  : : |        |
Db     371 RCWPSDSADDGWSPWSEWTSCSTSCGNGIQQRGRSCDSLNNRCEGSSVQTRTCHIQECDK 430

Qy     163 TAPRDTGAFDAVGEVEAWHRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLC 222
             :| |          |     :: |||| || :|| || |||             |||
Db     431 RFKQDGG----------W-----SHWSPWSSCSVTCGDGVITRI-------------RLC 462

Qy     223 NL-------RPCDVDIHTLIKAGKK--C-----LAVYQPEASMNFTLAGCISTRS----- 263
           |        :||: :     || ||  |       : |    : |  | : ||
Db     463 NSPSPQMNGKPCEGEARE-TKACKKDACPINGGWGPWSPWDICSVTCGGGVQKRSRLCNN 521
```

```
Qy    264 ----YQPKYC-GVCMDNRCCIPYKSKTIDVSFQCP-DGLGFSRQVLWINACFCNLSCRNP 317
              :  | | |   :|: |    :|       || ||          | ||  : |
Db    522 PTPQFGGKDCVGDVTENQIC----NKQ-----DCPIDG-------CLSNPCFAGVKC--- 562

Qy    318 NDIFADLESYPDFS 331
                  |||| |
Db    563 -------TSYPDGS 569


RESULT  15
T26972
hypothetical protein Y47H9C.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C;Accession: T26972
R;Harris, B.
submitted to the EMBL Data Library, October 1998
A;Reference number: Z20293
A;Accession: T26972
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1111 <WIL>
A;Cross-references: EMBL:AL032657; PIDN:CAA21739.1; GSPDB:GN00019; CESP:Y47H9C.4
A;Experimental source: clone Y47H9C
C;Genetics:
A;Gene: CESP:Y47H9C.4
A;Map position: 1
A;Introns: 50/2; 84/2; 150/1; 238/3; 342/3; 797/1; 851/1; 947/2; 1017/1; 1083/1
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology


  Query Match             7.3%;  Score 142;  DB 2;  Length 1111;
  Best Local Similarity   23.9%;  Pred. No. 0.0044;
  Matches   73;  Conservative   24;  Mismatches  119;  Indels   90;  Gaps   19;

Qy     23 CPPSPPRCPLGVSLITDGCE--CCKMCAQQLGDNCTEAAICDPHRGLYCDYSGDRPRYAI 80
          | |:   |        | ||  | |      | || : |:|| :    |:     |
Db    394 CAPNTGFCRCKPGFYGDNCELACSK---DSYGPNCEKQAMCDWNHASECN-----PETGS 445

Qy     81 GVC-----AQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLCLRVRPPRLWC 135
           ||      :     | ||         : ||| : | |    |||     : :  | |
Db    446 CVCKPGRTGKNCSEPCPLD--------FYGPNCAHQCQCNQRGVGCDGADGKCQCDRGW- 496

Qy    136 PHPRRVSIPGHCCEQWV--------CEDDAKRPRKTAPRDTGAFDAVGEVEAWHRNCIAY 187
                 || ||           ||   |      |:  |     ||       |
Db    497 --------TGHRCEHHCPADTFGANCEKRCK-----CPKGIGCDPITGE-------C--- 533

Qy    188 TSP----WSPCSTSC-----GLGVSTRISNVNAQCWPEQESRLCNLRP--CDVDIHTLI- 235
          | |     : |   |     | |       || :|  ::|:  |  :|     |  |
Db    534 TCPAGLQGANCDIGCPEGSYGPGCKLHCKCVNGKC--DKETGECTCQPGFFGSDCSTTCS 591

Qy    236 --KAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC- 292
            | |: |      | |   : | |         |  |     :|  |  :| :    :|
Db    592 KGKYGESC------ELSCPCSDASC-------SKQTG-----KCLCPLGTKGVSCDQKCD 633

Qy    293 PDGLGF 298
          |:  ||
Db    634 PNTFGF 639


Search completed: July 26, 2001, 08:36:29
Job time: 936 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 26, 2001, 08:36:08 ; Search time 12.6 Seconds
(without alignments)
910.761 Million cell updates/sec

Title: US-09-325-019-2
Perfect score: 1937
Sequence: 1 DFTPAPLEDTSSRPQFCKWP.....NPNDIFADLESYPDFSEIAN 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	869.5	44.9	349	1	CTGF_HUMAN	P29279 homo sapien
2	839.5	43.3	349	1	CTGF_PIG	O19113 sus scrofa
3	831.5	42.9	348	1	CTGF_MOUSE	P29268 mus musculu
4	793.5	41.0	349	1	CTGF_BOVIN	O18739 bos taurus
5	772.5	39.9	379	1	CYR6_MOUSE	P18406 mus musculu
6	764.5	39.5	381	1	CYR6_HUMAN	O00622 homo sapien
7	758	39.1	375	1	CE10_CHICK	P19336 gallus gall
8	743.5	38.4	351	1	NOV_CHICK	P28686 gallus gall
9	741.5	38.3	353	1	NOV_COTJA	P42642 coturnix co
10	726.5	37.5	354	1	NOV_MOUSE	Q64299 mus musculu
11	726.5	37.5	357	1	NOV_HUMAN	P48745 homo sapien
12	703	36.3	343	1	NOV_XENLA	P51609 xenopus lae
13	159	8.2	2482	1	VWF_PIG	Q28833 sus scrofa
14	151.5	7.8	5179	1	MUC2_HUMAN	Q02817 homo sapien
15	150	7.7	2531	1	NTC1_MOUSE	Q01705 mus musculu
16	150	7.7	2813	1	VWF_CANFA	Q28295 canis famil
17	150	7.7	2813	1	VWF_HUMAN	P04275 homo sapien
18	147.5	7.6	1178	1	TSP2_CHICK	P35440 gallus gall
19	144.5	7.5	837	1	MUCL_RAT	P98089 rattus norv
20	144	7.4	1173	1	TSP1_XENLA	P35448 xenopus lae
21	142.5	7.4	1170	1	TSP1_HUMAN	P07996 homo sapien
22	140.5	7.3	1170	1	TSP1_MOUSE	P35441 mus musculu
23	139	7.2	1700	1	BAR3_CHITE	Q03376 chironomus
24	138.5	7.2	2139	1	CRB_DROME	P10040 drosophila
25	138	7.1	555	1	DP87_DICDI	Q04503 dictyosteli
26	138	7.1	1056	1	MUC5_HUMAN	P98088 homo sapien
27	138	7.1	4289	1	TENX_HUMAN	P22105 homo sapien
28	137.5	7.1	810	1	NEL1_RAT	Q62919 rattus norv
29	136.5	7.0	1170	1	TSP1_BOVIN	Q28178 bos taurus
30	136	7.0	2437	1	NOTC_BRARE	P46530 brachydanio
31	135	7.0	2531	1	NTC1_RAT	Q07008 rattus norv
32	134.5	6.9	854	1	LDLR_CRIGR	P35950 cricetulus
33	134.5	6.9	864	1	LDLR_MOUSE	P35951 mus musculu
34	133.5	6.9	2444	1	NTC1_HUMAN	P46531 homo sapien
35	133	6.9	2524	1	NOTC_XENLA	P21783 xenopus lae
36	132	6.8	588	1	GRN_RAT	P23785 r granulins
37	131.5	6.8	1964	1	NTC4_MOUSE	P31695 mus musculu
38	130.5	6.7	1170	1	TSP2_BOVIN	Q95116 bos taurus
39	129.5	6.7	1408	1	SERR_DROME	P18168 drosophila
40	129	6.7	1172	1	TSP2_HUMAN	P35442 homo sapien
41	128.5	6.6	2911	1	FBN2_HUMAN	P35556 homo sapien
42	128.5	6.6	4544	1	LRP1_HUMAN	Q07954 homo sapien
43	128	6.6	3075	1	LMA1_HUMAN	P25391 homo sapien
44	126.5	6.5	1895	1	YLK3_CAEEL	P41951 caenorhabdi
45	126.5	6.5	2703	1	NOTC_DROME	P07207 drosophila

ALIGNMENTS

```
RESULT  1
CTGF_HUMAN
ID   CTGF_HUMAN     STANDARD;      PRT;   349 AA.
AC   P29279;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR.
GN   CTGF.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Umbilical vein endothelial cells;
RX   MEDLINE=91373462; PubMed=1654338;
RA   Bradham D.M., Igarashi A., Potter R.L., Grotendorst G.R.;
RT   "Connective tissue growth factor: a cysteine-rich mitogen secreted by
RT   human vascular endothelial cells is related to the SRC-induced
RT   immediate early gene product CEF-10.";
RL   J. Cell Biol. 114:1285-1294(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Umbilical vein endothelial cells;
RX   MEDLINE=93187114; PubMed=1293144;
RA   Igarashi A., Bradham D.M., Okochi H., Grotendorst G.R.;
RT   "Connective tissue growth factor.";
RL   J. Dermatol. 19:642-643(1992).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97207446; PubMed=9054739;
RA   Oemar B.S., Werner A., Garnier J.M., Do D.D., Godoy N., Nauck M.,
RA   Marz W., Rupp J., Pech M., Luescher T.F.;
RT   "Human connective tissue growth factor is expressed in advanced
RT   atherosclerotic lesions.";
RL   Circulation 95:831-839(1997).
CC   -!- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY
CC       HUMAN VASCULAR ENDOTHELIAL CELLS. THIS IMMEDIATE-EARLY PROTEIN
CC       MAY BIND ONE OF THE PDGF CELL SURFACE RECEPTORS.
CC   -!- SUBUNIT: MONOMER.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC       SHORT FORM; SEEM TO BE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
CC       PROTEIN FAMILY. CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 VWFC DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; M92934; AAA91279.1; -.
DR   EMBL; X78947; CAA55544.1; -.
DR   PIR; A40551; A40551.
DR   PIR; S44205; S44205.
DR   MIM; 121009; -.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001007; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00093; vwc; 1.
DR   PROSITE; PS00222; IGF_BINDING; 1.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS01208; VWFC; 1.
KW   Growth factor binding; Signal; Alternative splicing.
FT   SIGNAL       1     26       POTENTIAL.
FT   CHAIN       27    349       CONNECTIVE TISSUE GROWTH FACTOR.
FT   DOMAIN     101    167       VWFC.
FT   DOMAIN     256    330       CTCK.
FT   DISULFID   256    293       BY SIMILARITY.
FT   DISULFID   273    307       BY SIMILARITY.
FT   DISULFID   284    323       BY SIMILARITY.
FT   DISULFID   287    325       BY SIMILARITY.
FT   DISULFID   292    329       BY SIMILARITY.
FT   CARBOHYD    28     28       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   225    225       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   172    198       MISSING (IN SHORT ISOFORM).
SQ   SEQUENCE   349 AA;  38069 MW;  0ECF8470B357EA95 CRC64;


  Query Match             44.9%;  Score 869.5;  DB 1;  Length 349;
  Best Local Similarity   45.6%;  Pred. No. 4.5e-65;
  Matches  156;  Conservative   53;  Mismatches   98;  Indels   35;  Gaps    8;

Qy      12 SRP---QFCKWPCECPPSP-PRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
           |||   | |  || ||  | |||| ||||: ||| ||::||:|||: |||   ||||:||
Db      21 SRPAVGQNCSGPCRCPDEPAPRCPAGVSLVLDGCGCCRVCAKQLGELCTERDPCDPHKGL 80

Qy      68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
           :||: |      ||||    |  |:  |  | :|:||| :||| |||:|||||| ||| :
Db      81 FCDF-GSPANRKIGVCTAKDGAPCIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSM 139

Qy     127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
            || |   || |||| :|| |||:|||::         |:|      ||   | :|
Db     140 DVRLPSPDCPFPRRVKLPGKCCEEWVCDE---------PKDQ---TVVGPALAAYRLEDT 187

Qy     183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                     ||:  |: || || :||:|:|||::| || |  |::|||| :|||: |:
Db     188 FGPDPTMIRANCLVQTTEWSACSKTCGMGISTRVTNDNASCRLEKQSRLCMVRPCEADLE 247

Qy     233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
             || ||||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db     248 ENIKKGKKCIRTPKISKPIKFELSGCTSMKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKC 307

Qy     293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
           |||    : :::|  | |: :|   ||||  |   : | | :
Db     308 PDGEVMKKNMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 349


RESULT   2
CTGF_PIG
ID   CTGF_PIG       STANDARD;      PRT;   349 AA.
AC   O19113;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR.
GN   CTGF.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Uterus;
RX   MEDLINE=97390475; PubMed=9242708;
RA   Brigstock D.R., Steffen C.L., Kim G.Y., Vegunta R.K., Diehl J.R.,
RA   Harding P.A.;
RT   "Purification and characterization of novel heparin-binding growth
RT   factors in uterine secretory fluids. Identification as heparin-
RT   regulated Mr 10,000 forms of connective tissue growth factor.";
RL   J. Biol. Chem. 272:20275-20282(1997).
CC   -!- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY
CC       HUMAN VASCULAR ENDOTHELIAL CELLS. THIS IMMEDIATE-EARLY PROTEIN
CC       MAY BIND ONE OF THE PDGF CELL SURFACE RECEPTORS.
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
CC       PROTEIN FAMILY. CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 VWFC DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U83916; AAC48756.1; -.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001007; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00093; vwc; 1.
DR   PROSITE; PS00222; IGF_BINDING; 1.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS01208; VWFC; 1.
KW   Growth factor binding; Signal.
FT   SIGNAL       1     26       POTENTIAL.
FT   CHAIN       27    349       CONNECTIVE TISSUE GROWTH FACTOR.
FT   DOMAIN     101    167       VWFC.
FT   DOMAIN     256    330       CTCK.
FT   DISULFID   256    293       BY SIMILARITY.
FT   DISULFID   273    307       BY SIMILARITY.
FT   DISULFID   284    323       BY SIMILARITY.
FT   DISULFID   287    325       BY SIMILARITY.
FT   DISULFID   292    329       BY SIMILARITY.
SQ   SEQUENCE   349 AA;  38007 MW;  BB510E2B2B52D4A0 CRC64;


  Query Match             43.3%;  Score 839.5;  DB 1;  Length 349;
  Best Local Similarity   44.7%;  Pred. No. 1.4e-62;
  Matches  153;  Conservative   53;  Mismatches  101;  Indels   35;  Gaps    8;

Qy      12 SRP---QFCKWPCECPPSPPR-CPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
           |||   | |  |:|     | || ||||: ||| ||::||:|||: |||   ||||:||
Db      21 SRPASGQDCSGQCQCAAGKRRACPAGVSLVLDGCGCCRLCAKQLGELCTERDPCDPHKGL 80

Qy      68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
           :||: |      ||||    |  ||  |  | :|:||| :||| |||:|||||| ||| :
Db      81 FCDF-GSPANRKIGVCTAKDGAPCVFGGTVYRSGESFQSSCKYQCTCLDGAVGCVPLCSM 139

Qy     127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
            || |   || |||| :|| |||:|||::         |:|      ||   | :|
Db     140 DVRLPSPDCPFPRRVKLPGKCCEEWVCDE---------PKD---HTVVGPALAAYRLEDT 187

Qy     183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                     ||:  |: || || :||:|:|||::| || |  |::|||| :|||: |:
```

```
Db    188 FGPDPTMMRANCLVQTTEWSACSKTCGMGISTRVTNDNAFCRLEKQSRLCMVRPCEADLE 247

Qy    233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
          || |||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db    248 ENIKKGKKCIRTPKISKPVKFELSGCTSVKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKC 307

Qy    293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
          |||    : :::|  | |: :|   ||||  |   : | | :
Db    308 PDGEVMKKSMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 349


RESULT   3
CTGF_MOUSE
ID   CTGF_MOUSE     STANDARD;      PRT;   348 AA.
AC   P29268;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR (CTGF) (FISP-12 PROTEIN).
GN   CTGF OR FISP12 OR FISP-12.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91363290; PubMed=1888698;
RA   Ryseck R.-P., Macdonald-Bravo H., Mattei M.-G., Bravo R.;
RT   "Structure, mapping, and expression of fisp-12, a growth factor-
RT   inducible gene encoding a secreted cysteine-rich protein.";
RL   Cell Growth Differ. 2:225-233(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91229699; PubMed=2029337;
RA   Brunner A., Chinn J., Neubauer M.G., Purchio A.F.;
RT   "Identification of a gene family regulated by transforming growth
RT   factor-beta.";
RL   DNA Cell Biol. 10:293-300(1991).
CC   -!- TISSUE SPECIFICITY: TESTIS, SPLEEN, KIDNEY, LUNG, HEART, AND BRAIN
CC       (LOWEST LEVEL IN TESTIS AND HIGHEST IN LUNG).
CC   -!- INDUCTION: BY GROWTH FACTORS.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
CC       PROTEIN FAMILY. CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 VWFC DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M70641; AAA37627.1; -.
DR   EMBL; M70642; AAA37628.1; -.
DR   EMBL; M80263; AAA73135.1; -.
DR   PIR; A53228; A53228.
DR   MGD; MGI:95537; Fisp12.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001007; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00093; vwc; 1.
DR   PROSITE; PS00222; IGF_BINDING; 1.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS01208; VWFC; 1.
KW   Growth factor binding; Signal.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    348       CONNECTIVE TISSUE GROWTH FACTOR.
FT   DOMAIN      100    166       VWFC.
FT   DOMAIN      255    329       CTCK.
FT   DISULFID    255    292       BY SIMILARITY.
FT   DISULFID    272    306       BY SIMILARITY.
FT   DISULFID    283    322       BY SIMILARITY.
FT   DISULFID    286    324       BY SIMILARITY.
FT   DISULFID    291    328       BY SIMILARITY.
FT   CONFLICT    161    161       K -> E (IN REF. 2).
SQ   SEQUENCE   348 AA;  37793 MW;  735B65B6A711686F CRC64;


  Query Match             42.9%;  Score 831.5;  DB 1;  Length 348;
  Best Local Similarity   44.2%;  Pred. No. 6.3e-62;
  Matches  151;  Conservative   56;  Mismatches  100;  Indels   35;  Gaps    8;

Qy     12 SRP---QFCKWPCEC-PPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
          :||   | |   |:|   : | || ||||: ||| ||::||:|||: |||   ||||:||
Db     20 TRPATGQDCSAQCQCAAEAAPHCPAGVSLVLDGCGCCRVCAKQLGELCTERDPCDPHKGL 79

Qy     68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
          :||: |      ||||    |  ||  |  | :|:||| :||| |||:|||||| ||| :
Db     80 FCDF-GSPANRKIGVCTAKDGAPCVFGGSVYRSGESFQSSCKYQCTCLDGAVGCVPLCSM 138

Qy    127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
           || |   || |||| :|| ||::|||::          |:|     |||   | :|
Db    139 DVRLPSPDCPFPRRVKLPGKCCKEWVCDE---------PKDR---TAVGPALAAYRLEDT 186

Qy    183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                    ||:  |: || || :||:|:|||::| |  |  |::|||| :|||: |:
Db    187 FGPDPTMMRANCLVQTTEWSACSKTCGMGISTRVTNDNTFCRLEKQSRLCMVRPCEADLE 246

Qy    233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
          || |||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db    247 ENIKKGKKCIRTPKIAKPVKFELSGCTSVKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKC 306

Qy    293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
          |||    : :::|  | |: :|   ||||  |   : | | :
Db    307 PDGEIMKKNMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 348


RESULT   4
CTGF_BOVIN
ID   CTGF_BOVIN     STANDARD;      PRT;   349 AA.
AC   O18739;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR.
GN   CTGF.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Aorta;
RA   Liliensiek B., Lin Z., Fotsis T., Schimanski M., Bierhaus A.,
RA   Kanitz M., Kauffmann G., Schweigerer L., Ziegler R., Nawroth P.P.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY
CC       HUMAN VASCULAR ENDOTHELIAL CELLS. THIS IMMEDIATE-EARLY PROTEIN
CC       MAY BIND ONE OF THE PDGF CELL SURFACE RECEPTORS.
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
CC       PROTEIN FAMILY. CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 VWFC DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF000137; AAB66596.1; -.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001007; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00093; vwc; 1.
DR   PROSITE; PS00222; IGF_BINDING; FALSE_NEG.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS01208; VWFC; FALSE_NEG.
KW   Growth factor binding; Signal.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27    349       CONNECTIVE TISSUE GROWTH FACTOR.
FT   DOMAIN      101    167       VWFC.
FT   DOMAIN      256    330       CTCK.
FT   DISULFID    256    293       BY SIMILARITY.
FT   DISULFID    273    307       BY SIMILARITY.
FT   DISULFID    287    325       BY SIMILARITY.
FT   DISULFID    292    329       BY SIMILARITY.
SQ   SEQUENCE   349 AA;  38152 MW;  D919023AE40D212E CRC64;


  Query Match             41.0%;  Score 793.5;  DB 1;  Length 349;
  Best Local Similarity   43.1%;  Pred. No. 8.8e-59;
  Matches  148;  Conservative   52;  Mismatches  106;  Indels   37;  Gaps    8;

Qy      12 SRP----QFCKWPCECPPSP-PRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRG 66
           |||      |  ||:||  | |||| ||||: ||| || :||:|| : |||   ||||:|
Db      21 SRPASSQDCCSAPCQCPAGPAPRCPAGVSLVLDGCGCC-VCAKQLSELCTERDPCDPHKG 79

Qy      67 LYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC- 125
           |:||: |      ||||    |   :  |  | :|:||| :||| |||:||:||| |||
Db      80 LFCDF-GSPTNRKIGVCTAKDGAPYIFGGTVYQSGESFQSSCKYQCTCLDGSVGCVPLCS 138

Qy     126 LRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR--- 182
           : || |   || |||| :|| |||:||  |: :              ||   | :|
Db     139 VDVRLPSPDCPFPRRVKLPGKCCEEWVSRDEKEH-----------TVVGPALAAYRLED 186

Qy     183 -----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDI 231
                     ||   |: ||  | :||:|:|||::| || |  |::|||| :|||: |:
Db     187 TFGPDPTMIRANCQVQTTEWSAYSKTCGMGISTRVTNDNAFCRLEKQSRLCMVRPCEADL 246

Qy     232 HTLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQ 291
             || ||||:   :    : | |:|| | ::|: |: ||| | ||| |::: |: | |:
Db     247 EENIKKGKKCIRTPKISKPIKFQLSGCTSMKTYRAKFFGVCTDGRCCTPHRTTTLPVEFK 306

Qy     292 CPDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
           ||||    : :::|  | |: :|   ||||  |   : | | :
Db     307 CPDGEVMKKSMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 349


RESULT   5
CYR6_MOUSE
ID   CYR6_MOUSE     STANDARD;      PRT;   379 AA.
AC   P18406;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   CYR61 PROTEIN PRECURSOR (3CH61).
GN   CYR61 OR IGFBP10.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=Fibroblast;
RX   MEDLINE=90287146; PubMed=2355916;
RA   O'Brien T.P., Yang G.P., Sanders L., Lau L.F.;
RT   "Expression of cyr61, a growth factor-inducible immediate-early
RT   gene.";
RL   Mol. Cell. Biol. 10:3569-3577(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AJ; TISSUE=Embryonic fibroblast;
RX   MEDLINE=91288203; PubMed=2062642;
RA   Latinkic B.V., O'Brien T.P., Lau L.F.;
RT   "Promoter function and structure of the growth factor-inducible
RT   immediate early gene cyr61.";
RL   Nucleic Acids Res. 19:3261-3267(1991).
CC   -!- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
CC       PROTEINS; PROMOTES PROLIFERATION, MIGRATION AND ADHESION.
CC   -!- TISSUE SPECIFICITY: LOW IN KIDNEY, ADRENAL GLAND, TESTES, BRAIN,
CC       AND OVARY, MODERATE IN HEART, UTERUS, AND SKELETAL MUSCLE, HIGHEST
CC       IN LUNG.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED FROM G(0)/G(1) THROUGH MID-G(1) IN
CC       NORMAL CELLS, AND AT A CONSTANT LEVEL IN RAPIDLY GROWING CELLS.
CC   -!- INDUCTION: BY GROWTH FACTORS.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
CC       PROTEIN FAMILY. CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 VWFC DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M32490; AAA37512.1; -.
DR   EMBL; X56790; CAA40109.1; -.
DR   PIR; A35669; A35669.
DR   MGD; MGI:88613; Cyr61.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001007; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00093; vwc; 1.
DR   PROSITE; PS00222; IGF_BINDING; 1.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS01208; VWFC; 1.
KW   Growth factor binding; Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    379       CYR61 PROTEIN.
FT   DOMAIN       98    164       VWFC.
FT   DOMAIN      284    358       CTCK.
FT   DISULFID    284    321       BY SIMILARITY.
FT   DISULFID    301    335       BY SIMILARITY.
FT   DISULFID    312    351       BY SIMILARITY.
FT   DISULFID    315    353       BY SIMILARITY.
FT   DISULFID    320    357       BY SIMILARITY.
SQ   SEQUENCE   379 AA;  41709 MW;  FA6B5014B56A8EE9 CRC64;


  Query Match             39.9%;  Score 772.5;  DB 1;  Length 379;
  Best Local Similarity   39.6%;  Pred. No. 5.2e-57;
  Matches  141;  Conservative   56;  Mismatches  114;  Indels   45;  Gaps    6;

Qy      17 CKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGLYCDYSGDRP 76
           |   | ||   |:|  || |: ||| |||:||:|| ::|::   ||  :|| |:: |
Db      26 CPAACHCPLEAPKCAPGVGLVRDGCGCCKVCAKQLNEDCSKTQPCDHTKGLECNF-GASS 84
```

```
Qy      77 RYAIGVC-AQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-LRVRPPRLW 134
           |:| ||  |  |  :   | ||:|||||||: |||||||||| |||   :  | |
Db      85 TALKGICRAQSEGRPCEYNSRIYQNGESFQPNCKHQCTCIDGAVGCIPLCPQELSLPNLG 144

Qy     135 CPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHRN---------- 183
           ||:|| | : | |||:|||::|: :       |    ||  |||    |
Db     145 CPNPRLVKVSGQCCEEWVCDEDSIKDSLDDQDDLLGLDA-SEVELTRNNELIAIGKGSSL 203

Qy     184 ----------------------CIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESR 220
                                 ||  |: || || ||| |:|||::| | :|   :|:|
Db     204 KRLPVFGTEPRVLFNPLHAHGQKCIVQTTSWSQCSKSCGTGISTRVTNDNPECRLVKETR 263

Qy     221 LCNLRPCDVDIHTLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIP 280
           :| :|||   ::: :| ||||    :    : || ||| | : |:|||| |:| ||| |
Db     264 ICEVRPCGQPVYSSLKKGKKCSKTKKSPEPVRFTYAGCSSVKKYRPKYCGSCVDGRCCTP 323

Qy     281 YKSKTIDVSFQCPDGLGFSRQVLWINACFCNLSCRNPND-------IFADLESYPD 329
           :::|: : |:| ||  ||: |: | :| || :| :||:      :| |:  : |
Db     324 LQTRTVKMRFRCEDGEMFSKNVMMIQSCKCNYNCPHPNEASFRLYSLFNDIHKFRD 379


RESULT   6
CYR6_HUMAN
ID   CYR6_HUMAN     STANDARD;      PRT;   381 AA.
AC   O00622; O14934;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   CYR61 PROTEIN PRECURSOR (GIG1 PROTEIN) (INSULIN-LIKE GROWTH FACTOR-
DE   BINDING PROTEIN 10).
GN   CYR61 OR IGFBP10 OR GIG1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Albrecht C., von der Kammer H., Klaudiny J., Mayhaus M., Nitsch R.M.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97280750; PubMed=9135077;
RA   Jay P., Berge-Lefranc J.L., Marsollier C., Mejean C., Taviaux S.,
RA   Berta P.;
RT   "The human growth factor-inducible immediate early gene, CYR61, maps
RT   to chromosome 1p.";
RL   Oncogene 14:1753-1757(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RA   Kolesnikova T.V., Lau L.F.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Bi A.B., Yu L.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
CC       PROTEINS; PROMOTES PROLIFERATION, MIGRATION AND ADHESION (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
CC       PROTEIN FAMILY. CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 VWFC DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y12084; CAA72802.1; -.
DR   EMBL; U62015; AAB58319.1; -.
DR   EMBL; AF003594; AAB61240.1; -.
DR   EMBL; AF031385; AAB84227.1; -.
DR   MIM; 602369; -.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001007; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00093; vwc; 1.
DR   PROSITE; PS00222; IGF_BINDING; 1.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS01208; VWFC; 1.
KW   Growth factor binding; Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    381       CYR61 PROTEIN.
FT   DOMAIN       98    164       VWFC.
FT   DOMAIN      286    360       CTCK.
FT   DISULFID    286    323       BY SIMILARITY.
FT   DISULFID    303    337       BY SIMILARITY.
FT   DISULFID    314    353       BY SIMILARITY.
FT   DISULFID    317    355       BY SIMILARITY.
FT   DISULFID    322    359       BY SIMILARITY.
FT   CONFLICT    210    210       L -> I (IN REF. 4).
FT   CONFLICT    220    220       L -> R (IN REF. 4).
SQ   SEQUENCE   381 AA;  42026 MW;  FC0BD39C078CA0B1 CRC64;


  Query Match             39.5%;  Score 764.5;  DB 1;  Length 381;
  Best Local Similarity   39.7%;  Pred. No. 2.4e-56;
  Matches  142;  Conservative   57;  Mismatches  112;  Indels   47;  Gaps    8;

Qy      17 CKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGLYCDYSGDRP 76
           |   | ||   |:|  || |: ||| |||:||:|| ::|::   ||  :|| |:: |
Db      26 CPAACHCPLEAPKCAPGVGLVRDGCGCCKVCAKQLNEDCSKTQPCDHTKGLECNF-GASS 84

Qy      77 RYAIGVC-AQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-LRVRPPRLW 134
           |:| ||  |  |  :   | ||:|||||||: |||||||||| |||   :  | |
Db      85 TALKGICRAQSEGRPCEYNSRIYQNGESFQPNCKHQCTCIDGAVGCIPLCPQELSLPNLG 144

Qy     135 CPHPRRVSIPGHCCEQWVCEDDA-KRPRKTAPRDTG---AFDAVGEVEAWHRN------- 183
           ||:|| | : | |||:|||::|: | | :       |    |||  |||    |
Db     145 CPNPRLVKVTGQCCEEWVCDEDSIKDPMEDQDGLLGKELGFDA-SEVELTRNNELIAVGK 203

Qy     184 -------------------------CIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQE 218
                                    ||  |: || || :|| |:|||::| | :|   :|
Db     204 GSSLKRLPVFGMEPRILYNPLQGQKCIVQTTSWSQCSKTCGTGISTRVTNDNPECRLVKE 263

Qy     219 SRLCNLRPCDVDIHTLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCC 278
           :|:| :|||   ::: :| ||||    :    : || |||:| : |:|||| |:| |||
Db     264 TRICEVRPCGQPVYSSLKKGKKCSKTKKSPEPVRFTYAGCLSVKKYRPKYCGSCVDGRCC 323

Qy     279 IPYKSKTIDVSFQCPDGLGFSRQVLWINACFCNLSCRNPND-------IFADLESYPD 329
           |  ::|: : |:| ||  ||: |: | :| || :| : |:       :| |:  : |
Db     324 TPQLTRTVKMRFRCEDGETFSKNVMMIQSCKCNYNCPHANEAAFPFYRLFNDIHKFRD 381


RESULT   7
CE10_CHICK
ID   CE10_CHICK     STANDARD;      PRT;   375 AA.
AC   P19336;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   CEF-10 PROTEIN PRECURSOR.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89145206; PubMed=2537491;
RA   Simmons D.L., Levy D.B., Yannoni Y., Erikson R.L.;
RT   "Identification of a phorbol ester-repressible v-src-inducible gene.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:1178-1182(1989).
CC   -!- FUNCTION: PROBABLE SECRETED REGULATORY PROTEIN.
CC   -!- INDUCTION: BY V-SRC.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
CC       PROTEIN FAMILY. CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 VWFC DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J04496; AAA48661.1; -.
DR   PIR; A41428; A41428.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001007; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00093; vwc; 1.
DR   PROSITE; PS00222; IGF_BINDING; 1.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS01208; VWFC; 1.
KW   Growth factor binding; Signal.
FT   SIGNAL        1     22
FT   CHAIN        23    375       CEF-10 PROTEIN.
FT   DOMAIN       98    164       VWFC.
FT   DOMAIN      281    355       CTCK.
FT   DISULFID    281    318       BY SIMILARITY.
FT   DISULFID    298    332       BY SIMILARITY.
FT   DISULFID    309    348       BY SIMILARITY.
FT   DISULFID    312    350       BY SIMILARITY.
FT   DISULFID    317    354       BY SIMILARITY.
SQ   SEQUENCE   375 AA;  40651 MW;  95F28553BE35D5AE CRC64;


  Query Match             39.1%;  Score 758;  DB 1;  Length 375;
  Best Local Similarity   38.9%;  Pred. No. 8.1e-56;
  Matches  142;  Conservative   55;  Mismatches  100;  Indels   68;  Gaps    9;

Qy      17 CKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGLYCDYSGDRP 76
           |   |:|| : |:|  || |: ||| |||:||:|| ::|:    ||  :|| |:: |  |
Db      26 CPAVCQCPAAAPQCAPGVGLVPDGCGCCKVCAKQLNEDCSRTQPCDHTKGLECNF-GASP 84

Qy      77 RYAIGVC-AQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-LRVRPPRLW 134
             |:| ||  |  |  :   | ||:||||||: ||||||||| |||   :  | |
Db      85 AATNGICRAQSEGRPCEYNSKIYQNGESFQPNCKHQCTCIDGAVGCIPLCPQELSLPNLG 144

Qy     135 CPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWH------------ 181
           || || | :|| |||:||| |::|             ||: |:| :
Db     145 CPSPRLVKVPGQCCEEWVC-DESK------------DALEELEGFFSKEFGLDASEGEL 190

Qy     182 -RN----------------------------CIAYTSPWSPCSTSCGLGVSTRISNVN 210
            ||                           ||  |: || || :|| |:|||::| |
Db     191 TRNNELIAIVKGGLKMLPVFGSEPQSRAFENPKCIVQTTSWSQCSKTCGTGISTRVTNDN 250

Qy     211 AQCWPEQESRLCNLRPCDVDIHTLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCG 270
             |  :|:|:| :|||    :  :| ||||    :  : : || ||| | : |:|||||
Db     251 PDCKLIKETRICEVRPCGQPSYASLKKGKKCTKTKKSPSPVRFTYAGCSSVKKYRPKYCG 310

Qy     271 VCMDNRCCIPYKSKTIDVSFQCPDGLGFSRQVLWINACFCNLSCRNPNDIFADLESYPDF 330
           |:| ||| | :::|: : |:| ||  |:: |: | :| || :| : |      |:|| |
Db     311 SCVDGRCCTPQQTRTVKIRFRCDDGETFTKSVMMIQSCRCNYNCPHAN------EAYP-F 363

Qy     331 SEIAN 335
           : |
Db     364 YRLVN 368


RESULT   8
NOV_CHICK
ID   NOV_CHICK      STANDARD;      PRT;   351 AA.
AC   P28686;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   NOV PROTEIN PRECURSOR.
GN   NOV.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BROWN LEGHORN;
RX   MEDLINE=92107157; PubMed=1309586;
RA   Joliot V., Martinerie C., Dambrine G., Plassiart G., Brisac M.,
RA   Crochet J., Perbal B.;
RT   "Proviral rearrangements and overexpression of a new cellular gene
RT   (nov) in myeloblastosis-associated virus type 1-induced
RT   nephroblastomas.";
RL   Mol. Cell. Biol. 12:10-21(1992).
CC   -!- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
CC       GROWTH REGULATION. ITS OVEREXPRESSION IS ASSOCIATED WITH
CC       TUMORIGENESIS AND EXPRESSION OF A N-TERMINAL-TRUNCATED VERSION
CC       OF NOV GENE IN CHICKEN EMBRYONIC FIBROBLASTS (CEF) IS SUFFICIENT
CC       TO INDUCE THE TRANSFORMATION OF CEF IN VITRO.
CC   -!- TISSUE SPECIFICITY: BRAIN AND HEART, AND AT A LOWER LEVEL IN
CC       MUSCLE AND INTESTINE, IN THE EMBRYO. LUNG AND LESS SO IN BRAIN AND
CC       SPLEEN, IN ADULT CHICKEN.
CC   -!- DEVELOPMENTAL STAGE: MAV1-INDUCED NEPHROBLASTOMAS EXPRESS A HIGH
CC       LEVEL OF NOV GENE WHOSE TRANSCRIPTION IS NORMALLY ARRESTED IN
CC       ADULT KIDNEY.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
CC       PROTEIN FAMILY. CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 VWFC DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X59284; CAA41975.1; -.
DR   PIR; S20078; S20078.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001007; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00093; vwc; 1.
DR   PROSITE; PS00222; IGF_BINDING; 1.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS01208; VWFC; 1.
KW   Proto-oncogene; Growth factor binding; Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    351       NOV PROTEIN.
```

```
FT   DOMAIN      104    170       VWFC.
FT   DOMAIN      258    332       CTCK.
FT   DISULFID    258    295       BY SIMILARITY.
FT   DISULFID    275    309       BY SIMILARITY.
FT   DISULFID    286    325       BY SIMILARITY.
FT   DISULFID    289    327       BY SIMILARITY.
FT   DISULFID    294    331       BY SIMILARITY.
FT   CARBOHYD    274    274       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   351 AA;  38268 MW;  1ECB3FA3058C6797 CRC64;


  Query Match             38.4%;  Score 743.5;  DB 1;  Length 351;
  Best Local Similarity   41.2%;  Pred. No. 1.2e-54;
  Matches  135;  Conservative   49;  Mismatches  117;  Indels   27;  Gaps    7;

Qy       9 DTSSRPQFCKWPC--ECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRG 66
           : | |   |  ||   ||  ||||  ||  : ||| || :||:| |::|:    ||   |
Db      23 EVSGREAACPRPCGGRCPAEPPRCAPGVPAVLDGCGCCLVCARQRGESCSPLLPCDESGG 82

Qy      67 LYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC- 125
           ||||  |       |:|  : |  || ||: | ||::|||:||| ||| || :|| | |
Db      83 LYCD-RGPEDGGGAGICMVLEGDNCVFDGMIYRNGETFQPSCKYQCTCRDGQIGCLPRCN 141

Qy     126 LRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRD---TGAFDAVG------- 175
           | :  |   || ||:: :|| |||:|||:          |||    | |
Db     142 LGLLLPGPDCPFPRKIEVPGECCEKWVCD---------PRDEVLLGGFAMAAYRQEATL 191

Qy     176 --EVEAWHRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHT 233
             :|     |||  |: || || |||:| |||::| | ||   :::||| :|||: :
Db     192 GIDVSDSSANCIEQTTEWSACSKSCGMGFSTRVTNRNQQCEMVKQTRLCMMRPCENE-EP 250

Qy     234 LIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCP 293
           | ||||:   :   :: |    | | ::|:|:|||:| | ||| |: :||| | |:||
Db     251 SDKKGKKCIQTKKSMKAVRFEYKNCTSVQTYKPRYCGLCNDGRCCTPHNTKTIQVEFRCP 310

Qy     294 DGLGFSRQVLWINACFCNLSCRNPNDIF 321
           |    : :: || | |: :|   |: |
Db     311 QGKFLKKPMMLINTCVCHGNCPQSNNAF 338


RESULT   9
NOV_COTJA
ID   NOV_COTJA      STANDARD;      PRT;   353 AA.
AC   P42642;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   NOV PROTEIN PRECURSOR.
GN   NOV.
OS   Coturnix coturnix japonica (Japanese quail).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Coturnix.
OX   NCBI_TaxID=93934;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Weiskirchen R., Bister K.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
CC       GROWTH REGULATION (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
CC       PROTEIN FAMILY. CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 VWFC DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U13063; AAA21128.1; -.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001007; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00093; vwc; 1.
DR   PROSITE; PS00222; IGF_BINDING; 1.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS01208; VWFC; 1.
KW   Proto-oncogene; Growth factor binding; Signal.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27    353       NOV PROTEIN.
FT   DOMAIN      106    172       VWFC.
FT   DOMAIN      260    334       CTCK.
FT   DISULFID    260    297       BY SIMILARITY.
FT   DISULFID    277    311       BY SIMILARITY.
FT   DISULFID    288    327       BY SIMILARITY.
FT   DISULFID    291    329       BY SIMILARITY.
FT   DISULFID    296    333       BY SIMILARITY.
FT   CARBOHYD    276    276       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   353 AA;  38667 MW;  717D9F8533882E89 CRC64;


  Query Match             38.3%;  Score 741.5;  DB 1;  Length 353;
  Best Local Similarity   41.2%;  Pred. No. 1.8e-54;
  Matches  135;  Conservative   49;  Mismatches  117;  Indels   27;  Gaps    7;

Qy       9 DTSSRPQFCKWPC--ECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRG 66
           : : |   |  ||   ||  ||||  ||  : ||| || :||:| |::|:    ||   |
Db      25 EVNGREAPCPRPCGGRCPAEPPRCAPGVPAVLDGCGCCLVCARQRGESCSPLLPCDESGG 84

Qy      67 LYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC- 125
           ||||  |       |:|  : |  || ||: | ||::|||:||| ||| || :|| | |
Db      85 LYCD-RGPEDGGGTGICMVLEGDNCVFDGMIYRNGETFQPSCKYQCTCRDGQIGCLPRCN 143

Qy     126 LRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRD---TGAFDAVG------- 175
           | :  |   || ||:: :|| |||:||||          |||    | |
Db     144 LGLLLPGPDCPFPRKIEVPGECCEKWVCE---------PRDEVLLGGFAMAAYRQEATL 193

Qy     176 --EVEAWHRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHT 233
             :|     |||  |: || || |||:| |||::| | ||   :::||| :|||: :
Db     194 GIDVSDSSANCIEQTTEWSACSRSCGMGFSTRVTNRNQQCEMVKQTRLCMMRPCENE-EP 252

Qy     234 LIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCP 293
           | ||||:   :   :: |    | | ::|:|:|||:| | ||| |: :||| | |:||
Db     253 SDKKGKKCIRTKKSMKAVRFEYKNCTSVQTYKPRYCGLCNDGRCCTPHNTKTIQVEFRCP 312

Qy     294 DGLGFSRQVLWINACFCNLSCRNPNDIF 321
           |    : :: || | |: :|   |: |
Db     313 QGKFLKKPMMLINTCVCHGNCPQSNNAF 340


RESULT  10
NOV_MOUSE
ID   NOV_MOUSE      STANDARD;      PRT;   354 AA.
AC   Q64299;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   NOV PROTEIN HOMOLOG PRECURSOR (NOVH).
GN   NOV.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SV, AND ICR; TISSUE=Brain;
RX   MEDLINE=97131523; PubMed=8975721;
```

```
RA   Snaith M.R., Natarajan D., Taylor L.B., Choi C.P., Martinerie C.,
RA   Perbal B., Schofield P.N., Boulter C.A.;
RT   "Genomic structure and chromosomal mapping of the mouse nov gene.";
RL   Genomics 38:425-428(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6;
RX   MEDLINE=96204003; PubMed=8622864;
RA   Martinerie C., Chevalier G., Rauscher F.J. III, Perbal B.;
RT   "Regulation of nov by WT1: a potential role for nov in
RT   nephrogenesis.";
RL   Oncogene 12:1479-1492(1996).
CC   -!- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
CC       GROWTH REGULATION (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
CC       PROTEIN FAMILY. CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 VWFC DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X97863; CAA66457.1; -.
DR   EMBL; Y09257; CAA70454.1; -.
DR   EMBL; X96585; CAA65404.1; -.
DR   MGD; MGI:109185; Nov.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001007; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00093; vwc; 1.
DR   PROSITE; PS00222; IGF_BINDING; 1.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS01208; VWFC; 1.
KW   Proto-oncogene; Growth factor binding; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    354       NOV PROTEIN HOMOLOG.
FT   DOMAIN      102    168       VWFC.
FT   DOMAIN      261    335       CTCK.
FT   DISULFID    261    298       BY SIMILARITY.
FT   DISULFID    278    312       BY SIMILARITY.
FT   DISULFID    289    328       BY SIMILARITY.
FT   DISULFID    292    330       BY SIMILARITY.
FT   DISULFID    297    334       BY SIMILARITY.
FT   CARBOHYD     91     91       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    277    277       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   354 AA;  38928 MW;  08ECE8CFC67829DE CRC64;


  Query Match             37.5%;  Score 726.5;  DB 1;  Length 354;
  Best Local Similarity   42.6%;  Pred. No. 3.1e-53;
  Matches  144;  Conservative   41;  Mismatches  122;  Indels   31;  Gaps    8;

Qy      14 PQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGLYCDYSG 73
           |  |  | :||   | |  ||  : ||| || :||:| |::|:|   ||   ||||| |
Db      30 PSRC--PPKCPSISPTCAPGVRSVLDGCSCCPVCARQRGESCSEMRPCDQSSGLYCDRSA 87

Qy      74 DRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-LRVRPPR 132
           | |    |:|    |  || ||| | ||: |:|||:| ||| || :|| | | | |  |
Db      88 D-PNNQTGICMVPEGDNCVFDGVIYRNGEKFEPNCQYFCTCRDGQIGCLPRCQLDVLLPG 146

Qy     133 LWCPHPRRVSIPGHCCEQWVCEDD----------AKRPRKTAPRDTGAFDAVG-EV 177
             || ||:|::|| |||:| |  |              | ||  |          || ||
Db     147 PDCPAPRKVAVPGECCEKWTCGSDEQGTQGTLGGLALPAYRPEAT----------VGVEV 196

Qy     178 EAWHRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLI-K 236
             |||  |: || || |||:|||||::| | ||   :::||| :|||: :   :  |
Db     197 SDSSINCIEQTTEWSACSKSCGMGVSTRVTNRNRQCEMVKQTRLCIVRPCEQEPEEVTDK 256

Qy     237 AGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGL 296
           |||||   :  :::    | |  :|:|::|||| | ||| |: :||| | |||  |
Db     257 KGKKCLRTKKSLKAIHLQFENCTSLYTYKPRFCGVCSDGRCCTPHNTKTIQVEFQCLPGE 316

Qy     297 GFSRQVLWINACFCNLSCRNPNDIF-ADLESYPDFSEI 333
             : |: |  | |  :|   |: |  |||      ||
Db     317 IIKKPVMVIGTCTCYSNCPQNNEAFLQDLELKTSRGEI 354


RESULT  11
NOV_HUMAN
ID   NOV_HUMAN      STANDARD;      PRT;   357 AA.
AC   P48745;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   NOV PROTEIN HOMOLOG PRECURSOR (NOVH).
GN   NOV.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=94336229; PubMed=7520150;
RA   Martinerie C., Huff V., Joubert I., Badzioch M., Saunders G.,
RA   Strong L., Perbal B.;
RT   "Structural analysis of the human nov proto-oncogene and expression
RT   in Wilms tumor.";
RL   Oncogene 9:2729-2732(1994).
CC   -!- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
CC       GROWTH REGULATION (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: INCREASED EXPRESSION IN WILMS TUMOR OF THE
CC       STROMAL TYPE.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
CC       PROTEIN FAMILY. CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 VWFC DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X78351; CAA55146.1; -.
DR   EMBL; X78352; CAA55146.1; JOINED.
DR   EMBL; X78353; CAA55146.1; JOINED.
DR   EMBL; X78354; CAA55146.1; JOINED.
DR   EMBL; X96584; CAA65403.1; -.
DR   MIM; 164958; -.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001007; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00093; vwc; 1.
DR   PROSITE; PS00222; IGF_BINDING; 1.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS01208; VWFC; 1.
KW   Proto-oncogene; Growth factor binding; Signal.
FT   SIGNAL        1     27       POTENTIAL.
```

```
FT   CHAIN        28    357       NOV PROTEIN HOMOLOG.
FT   DOMAIN      108    174       VWFC.
FT   DOMAIN      264    338       CTCK.
FT   DISULFID    264    301       BY SIMILARITY.
FT   DISULFID    281    315       BY SIMILARITY.
FT   DISULFID    292    331       BY SIMILARITY.
FT   DISULFID    295    333       BY SIMILARITY.
FT   DISULFID    300    337       BY SIMILARITY.
FT   CARBOHYD     97     97       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    280    280       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   357 AA;  39162 MW;  035D5BF4576BD85B CRC64;


  Query Match             37.5%;  Score 726.5;  DB 1;  Length 357;
  Best Local Similarity   42.0%;  Pred. No. 3.1e-53;
  Matches  137;  Conservative   46;  Mismatches  116;  Indels   27;  Gaps    8;

Qy      10 TSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGLYC 69
           |   | |  |  || :|| |  ||  : ||| || :||:| |::|::   ||   ||||
Db      32 TQRCPPQC--PGRCPATPPTCAPGVRAVLDGCSCCLVCARQRGESCSDLEPCDESSGLYC 89

Qy      70 DYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-LRV 128
           | | | |    |:|  | |  || ||| | :|: |||:||: ||| || :|| | | | |
Db      90 DRSAD-PSNQTGICTAVEGDNCVFDGVIYRSGEKFQPSCKFQCTCRDGQIGCVPRCQLDV 148

Qy     129 RPPRLWCPHPRRVSIPGHCCEQWVC---EDD--------AKRPRKTAPRDTGAFDAVG-E 176
             |   || ||:| :|| |||:|:|   |:|        | ||  |          :| |
Db     149 LLPEPNCPAPRKVEVPGECCEKWICGPDEEDSLGGLTLAAYRPEAT----------LGVE 198

Qy     177 VEAWHRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVD-IHTLI 235
           |     |||  |: |: || |||:| |||::| | ||   :::|||  :|||: :
Db     199 VSDSSVNCIEQTTEWTACSKSCGMGFSTRVTNRNRQCEMLKQTRLCMVRPCEQEPEQPTD 258

Qy     236 KAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDG 295
           | |||||   :   :::     | |  :|:|::|||| | ||| |: :|||   |||  |
Db     259 KKGKKCLRTKKSLKAIHLQFKNCTSLHTYKPRFCGVCSDGRCCTPHNTKTIQAEFQCSPG 318

Qy     296 LGFSRQVLWINACFCNLSCRNPNDIF 321
             : |: |  | |: :|   |: |
Db     319 QIVKKPVMVIGTCTCHTNCPKNNEAF 344


RESULT  12
NOV_XENLA
ID   NOV_XENLA      STANDARD;      PRT;   343 AA.
AC   P51609;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   NOV PROTEIN HOMOLOG PRECURSOR (XNOV).
GN   NOV.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96257227; PubMed=8666280;
RA   Ying Z., King M.L.;
RT   "Isolation and characterization of xnov, a Xenopus laevis ortholog of
RT   the chicken nov gene.";
RL   Gene 171:243-248(1996).
CC   -!- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
CC       GROWTH REGULATION (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
CC       PROTEIN FAMILY. CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN SUBFAMILY.
CC   -!- SIMILARITY: CONTAINS 1 VWFC DOMAIN.
CC   -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U37063; AAB17096.1; -.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001007; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00093; vwc; 1.
DR   PROSITE; PS00222; IGF_BINDING; 1.
DR   PROSITE; PS01208; VWFC; 1.
DR   PROSITE; PS01185; CTCK_1; FALSE_NEG.
DR   PROSITE; PS01225; CTCK_2; 1.
KW   Growth factor binding; Signal.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    343       NOV PROTEIN HOMOLOG.
FT   DOMAIN       93    159       VWFC.
FT   DOMAIN      249    323       CTCK.
FT   DISULFID    249    286       BY SIMILARITY.
FT   DISULFID    266    300       BY SIMILARITY.
FT   DISULFID    277    316       BY SIMILARITY.
FT   DISULFID    280    318       BY SIMILARITY.
FT   DISULFID    285    322       BY SIMILARITY.
FT   CARBOHYD    265    265       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   343 AA;  38070 MW;  677D7078EB21365F CRC64;


  Query Match             36.3%;  Score 703;  DB 1;  Length 343;
  Best Local Similarity   41.7%;  Pred. No. 2.7e-51;
  Matches  133;  Conservative   50;  Mismatches  110;  Indels   26;  Gaps    8;

Qy      15 QFCKWPC-ECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGLYCDYSG 73
           | |   | :||  || |   | || ||| || :||:| |::|:    |   :||||:::
Db      19 QKCPSQCDQCPEEPPSCAPSVLLILDGCGCCPVCARQEGESCSHLNPCQEDKGLYCEFNA 78

Qy      74 DRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-LRVRPPR 132
           | ||   | |  : |  || ||| | | :||||:|||:|||::| :|| | | | :  |
Db      79 D-PRMETGTCMALEGNSCVFDGVVYRNRESFQPSCKYHCTCLNGHIGCVPRCNLDLLLPG 137

Qy     133 LWCPHPRRVSIPGHCCEQWVCEDD----------AKRPRKTAPRDTGAFDAVGEVEAWH 181
             || |||| :|| |||:|||:              | ||  |        ||     |
Db     138 PDCPFPRRVKVPGECCEKWVCDSKEEMAIGGFAMAAYRPEATL-----GIDASDTSFA-- 190

Qy     182 RNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLI-KAGKK 240
             ||| |: || || :||:|||:|::| ||:|  ::: ||| :| |: :    : | |||
Db     191 --CIAQTTEWSACSKTCGMGVSSRVTNRNARCEMQKQIRLCMVRSCEEEPGWHVEKKGKK 248

Qy     241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 300
           |: | :    ::|    | | : |:||:|| | | ||| |: :||: | | ||      :
Db     249 CVRVRKTTKPIHFHYKNCTSVQPYKPKFCGQCSDGRCCTPHSTKTMHVEFVCPQKRIVKK 308

Qy     301 QVLWINACFCNLSCRNPND 319
             |: |: | |: :|  | |
Db     309 PVMVISTCVCHYNC--PQD 325


RESULT  13
VWF_PIG
ID   VWF_PIG        STANDARD;      PRT;  2482 AA.
AC   Q28833;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   VON WILLEBRAND FACTOR PRECURSOR (VWF) (FRAGMENT).
GN   F8VWF OR VWF.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
```

```
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Seaman W.T., Read M.S., Bellinger D.A., Nichols T.C.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 397-553 FROM N.A.
RX   MEDLINE=93356762; PubMed=8352759;
RA   Lavergne J.M., Piao Y.C., Ferreira V., Kerbiriou-Nabias D.,
RA   Bahnak B.R., Meyer D.;
RT   "Primary structure of the factor VIII binding domain of human, porcine
RT   and rabbit von Willebrand factor.";
RL   Biochem. Biophys. Res. Commun. 194:1019-1024(1993).
CC   -!- FUNCTION: IMPORTANT IN THE MAINTENANCE OF HOMEOSTASIS, IT
CC       PARTICIPATES IN PLATELET-VESSEL WALL INTERACTIONS BY FORMING A
CC       NONCOVALENT COMPLEX WITH COAGULATION FACTOR VIII AT THE SITE OF
CC       VASCULAR INJURY (BY SIMILARITY).
CC   -!- SUBUNIT: MULTIMERIC (BY SIMILARITY).
CC   -!- PTM: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR
CC       INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).
CC   -!- SIMILARITY: CONTAINS 3 VWFC DOMAINS.
CC   -!- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
CC   -!- SIMILARITY: SOME, TO SILKWORM HEMOCYTIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF052036; AAC06229.1; -.
DR   EMBL; S64541; AAB27829.2; -.
DR   HSSP; P04275; 1ATZ.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR001007; -.
DR   InterPro; IPR001846; -.
DR   InterPro; IPR002035; -.
DR   InterPro; IPR002919; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00092; vwa; 3.
DR   Pfam; PF00093; vwc; 3.
DR   Pfam; PF00094; vwd; 3.
DR   Pfam; PF01826; TIL; 3.
DR   PRINTS; PR00365; ENDOTHELIN.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   PROSITE; PS50234; VWFA; 3.
DR   PROSITE; PS01208; VWFC; 3.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
KW   Blood coagulation; Platelet; Glycoprotein; Extracellular matrix;
KW   Plasma; Endothelial cell; Repeat; Cell adhesion.
FT   NON_TER       1      1
FT   PROPEP       <1    437       BY SIMILARITY.
FT   CHAIN        438   2482       VON WILLEBRAND FACTOR.
FT   DOMAIN        62    215       VWFD 2.
FT   DOMAIN       438    461       AMINO-TERMINAL.
FT   DOMAIN       462    507       E1.
FT   DOMAIN       500    527       CX.
FT   DOMAIN       541    687       VWFD 3.
FT   DOMAIN       947   1127       VWFA 1.
FT   DOMAIN      1167   1334       VWFA 2.
FT   DOMAIN      1360   1540       VWFA 3.
FT   DOMAIN      1619   1771       VWFD 4.
FT   DOMAIN      1885   1930       E2.
FT   DOMAIN      1924   1997       VWFC 1.
FT   DOMAIN      2098   2164       VWFC 2.
FT   DOMAIN      2249   2319       VWFC 3.
FT   DOMAIN      2393   2481       CTCK.
FT   SITE        2176   2178       CELL ATTACHMENT SITE (POTENTIAL).
FT   DISULFID     441    482       BY SIMILARITY.
FT   DISULFID     450    478       BY SIMILARITY.
FT   DISULFID     484    495       BY SIMILARITY.
FT   DISULFID     541    670       BY SIMILARITY.
FT   DISULFID     563    705       BY SIMILARITY.
FT   DISULFID     572    667       BY SIMILARITY.
FT   DISULFID     588    595       BY SIMILARITY.
FT   DISULFID     734    758       BY SIMILARITY.
FT   DISULFID     745    785       BY SIMILARITY.
FT   DISULFID     763    765       BY SIMILARITY.
FT   DISULFID     827    839       BY SIMILARITY.
FT   DISULFID     823    843       BY SIMILARITY.
FT   DISULFID     800    804       BY SIMILARITY.
FT   DISULFID     870    873       BY SIMILARITY.
FT   DISULFID     908    911       BY SIMILARITY.
FT   DISULFID     942   1128       BY SIMILARITY.
FT   DISULFID    1338   1339       BY SIMILARITY.
FT   DISULFID    1355   1541       BY SIMILARITY.
FT   DISULFID    1548   1573       BY SIMILARITY.
FT   DISULFID    1568   1609       OR 1611 (BY SIMILARITY).
FT   DISULFID    1641   1792       BY SIMILARITY.
FT   DISULFID    1619   1754       BY SIMILARITY.
FT   DISULFID    1596   1757       BY SIMILARITY.
FT   DISULFID    1662   1670       BY SIMILARITY.
FT   DISULFID    2393   2443       BY SIMILARITY.
FT   DISULFID    2408   2457       BY SIMILARITY.
FT   DISULFID    2419   2473       BY SIMILARITY.
FT   DISULFID    2423   2475       BY SIMILARITY.
FT   DISULFID       ?   2480       BY SIMILARITY.
FT   CARBOHYD     340    340       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     905    905       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1184   1184       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1243   1243       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1640   1640       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1892   1892       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    1959   1959       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    2026   2026       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    2069   2069       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    2215   2215       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    2254   2254       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    2459   2459       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   2482 AA;  272394 MW;  D499B7DDFBBCAEDD CRC64;

  Query Match             8.2%;  Score 159;  DB 1;  Length 2482;
  Best Local Similarity   23.7%;  Pred. No. 1.6e-05;
  Matches   94;  Conservative   44;  Mismatches  120;  Indels  138;  Gaps   31;

Qy      14 PQFCKWPCECPPSPPRCPLGVSLITDGCE--CCKMC---AQQLGDNCTEAAICDPHR--- 65
           | ||      |||          ||: : ||  | : |   :   ||: :|   | ||:
Db    1865 PDFC--AVSCPP---------SLVYNHCEHGCPRHCEGNSSSCGDHPSEGCFCPPHQVML 1913

Qy      66 GLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPN---CKYNCTCIDG-AVGC 121
           | |       |  |   | | |      ||:|:   ::: |:   |:  |||: |  | |
Db    1914 GSSC-----VPEEA---CTQCVDD----DGIRHQFLETWVPDHQPCQI-CTCLSGRRVNC 1960

Qy     122 T---------PLC-----LRVRP------PRLWC-----------PH-----PRRVSIP 144
           |         | |      |:|       |   |            ||         :: |
Db    1961 TLQPCPTARAPACGLCEVARLRQEAHQCCPEYECVCDLVSCDLPPVPHCEGGLQPTLTNP 2020

Qy     145 GHCCEQWVC----EDDAKRPRKTAP-------RDTGAFDAVGEVEAWHRNCIAYTSPWSP 193
           | |   : |    |:  : |  : |       | |   |   | |    ||:  |
Db    2021 GECRPNFTCACRKEECPRGPLPSCPPHRTPALRKTQCCD---EYEC-ACNCVNTT----- 2071

Qy     194 CSTSCGLG-VSTRISN----VNAQCWPEQESRLCNLR-----------PCDVDIHTLIK 236
             || || ::: ::|       | |:   ::|  |              |||
Db    2072 --LSCPLGYLASTVTNDCGCTTTTCLPD---KVCVHRGTVYPVGQFWEEGCDV------- 2119

Qy     237 AGKKCLAVYQPEASMNFTLAGCIS---TRSYQPKYCGVCMDNRC---CIPYKSKTIDVSF 290
               |     :| |   :| |       | :| :  |  :  |   |:|   | :  ||
Db    2120 ----CTCTDLEDAVMGLRVAQCAQKPCEDSCRPGFTYVLHEGECCGKCLPSACKVVIGSF 2175

Qy     291 QCPDGLGFSRQV--LWI---NACFCNLSCRNPNDIF 321
           : | : : : |   |    | | |   |   ::|
Db    2176 R-GDSVSYWKSVGSHWASPENPCLINECVRVKEEVF 2210
```

```
RESULT  14
MUC2_HUMAN
ID   MUC2_HUMAN     STANDARD;      PRT;  5179 AA.
AC   Q02817; Q14878;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   MUCIN 2 PRECURSOR (INTESTINAL MUCIN 2).
GN   MUC2 OR SMUC.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Intestine;
RX   MEDLINE=94132002; PubMed=8300571;
RA   Gum J.R. Jr., Hicks J.W., Toribara N.W., Siddiki B., Kim Y.S.;
RT   "Molecular cloning of human intestinal mucin (MUC2) cDNA.
RT   Identification of the amino terminus and overall sequence similarity
RT   to prepro-von Willebrand factor.";
RL   J. Biol. Chem. 269:2440-2446(1994).
RN   [2]
RP   SEQUENCE OF 626-1895 AND 4196-5179 FROM N.A.
RC   TISSUE=Colon;
RX   MEDLINE=93016075; PubMed=1400449;
RA   Gum J.R. Jr., Hicks J.W., Toribara N.W., Rothe E.-M., Lagace R.E.,
RA   Kim Y.S.;
RT   "The human MUC2 intestinal mucin has cysteine-rich subdomains located
RT   both upstream and downstream of its central repetitive region.";
RL   J. Biol. Chem. 267:21375-21383(1992).
RN   [3]
RP   SEQUENCE OF 1343-1895 AND 4176-4195 FROM N.A.
RX   MEDLINE=91358717; PubMed=1885763;
RA   Toribara N.W., Gum J.R. Jr., Culhane P.J., Lagace R.E., Hicks J.W.,
RA   Petersen G.M., Kim Y.S.;
RT   "MUC-2 human small intestinal mucin gene structure. Repeated arrays
RT   and polymorphism.";
RL   J. Clin. Invest. 88:1005-1013(1991).
CC   -!- FUNCTION: COATS THE EPITHELIA OF THE INTESTINES, AIRWAYS, AND
CC       OTHER MUCUS MEMBRANE-CONTAINING ORGANS. THOUGHT TO PROVIDE A
CC       PROTECTIVE, LUBRICATING BARRIER AGAINST PARTICLES AND INFECTIOUS
CC       AGENTS AT MUCOSAL SURFACES.
CC   -!- SUBUNIT: MULTIMERIC.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- TISSUE SPECIFICITY: COLON, SMALL INTESTINE, COLONIC TUMORS,
CC       BRONCHUS, CERVIX AND GALL BLADDER.
CC   -!- PTM: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR
CC       INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).
CC   -!- POLYMORPHISM: THE NUMBER OF REPEATS IS HIGHLY POLYMORPHIC AND
CC       VARIES AMONG DIFFERENT ALLELES.
CC   -!- SIMILARITY: THE N-TERMINAL DOMAIN SHOWS SOME SIMILARITY TO THAT
CC       OF SILKWORM HEMOCYTIN.
CC   -!- SIMILARITY: CONTAINS 2 VWFC DOMAINS.
CC   -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L21998; AAB95295.1; -.
DR   EMBL; M74027; AAA59875.1; -.
DR   EMBL; M94131; AAA59163.1; -.
DR   EMBL; M94132; AAA59164.1; -.
DR   MIM; 158370; -.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000561; -.
DR   InterPro; IPR001007; -.
DR   InterPro; IPR001846; -.
DR   InterPro; IPR002400; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00094; vwd; 4.
DR   PRINTS; PR00438; GFCYSKNOT.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS01208; VWFC; 2.
KW   Glycoprotein; Repeat; Signal.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21   5179       MUCIN 2.
FT   DOMAIN     1401   1747       APPROXIMATE REPEATS.
FT   REPEAT     1401   1416       1.
FT   REPEAT     1417   1432       2.
FT   REPEAT     1433   1448       3.
FT   REPEAT     1449   1464       4.
FT   REPEAT     1465   1471       5.
FT   REPEAT     1472   1478       6.
FT   REPEAT     1479   1494       7A.
FT   REPEAT     1495   1517       7B.
FT   REPEAT     1518   1533       8A.
FT   REPEAT     1534   1556       8B.
FT   REPEAT     1557   1572       9A.
FT   REPEAT     1573   1596       9B.
FT   REPEAT     1597   1612       10A.
FT   REPEAT     1613   1635       10B.
FT   REPEAT     1636   1651       11A.
FT   REPEAT     1652   1675       11B.
FT   REPEAT     1676   1683       12.
FT   REPEAT     1684   1699       13.
FT   REPEAT     1700   1715       14.
FT   REPEAT     1716   1731       15.
FT   REPEAT     1732   1747       16.
FT   DOMAIN     4815   4886       VWFC 1.
FT   DOMAIN     4924   4991       VWFC 2.
FT   DOMAIN     5075   5160       CTCK.
FT   DISULFID   5075   5122       BY SIMILARITY.
FT   DISULFID   5089   5136       BY SIMILARITY.
FT   DISULFID   5098   5152       BY SIMILARITY.
FT   DISULFID   5102   5154       BY SIMILARITY.
FT   DISULFID      ?   5159       BY SIMILARITY.
FT   CARBOHYD    163    163       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    423    423       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    670    670       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    770    770       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    894    894       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1139   1139       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1154   1154       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1215   1215       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1230   1230       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1246   1246       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1787   1787       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1820   1820       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4339   4339       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4351   4351       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4362   4362       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4373   4373       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4422   4422       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4438   4438       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4502   4502       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4616   4616       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4627   4627       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4752   4752       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4787   4787       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4881   4881       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4888   4888       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4955   4955       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   4970   4970       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   5019   5019       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   5038   5038       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   5069   5069       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   CONFLICT   1351   1351       H -> L (IN REF. 3).
FT   CONFLICT   1412   1412       T -> S (IN REF. 3).
FT   CONFLICT   1449   1449       L -> P (IN REF. 3).
FT   CONFLICT   1504   1504       M -> T (IN REF. 3).
FT   CONFLICT   4192   4192       G -> S (IN REF. 2).
SQ   SEQUENCE   5179 AA;  540295 MW;  85CD7571FB9A5663 CRC64;


  Query Match             7.8%;  Score 151.5;  DB 1;  Length 5179;
  Best Local Similarity   20.5%;  Pred. No. 0.00013;
  Matches   99;  Conservative   40;  Mismatches  156;  Indels  189;  Gaps   26;

Qy       1 DFTPAPLEDTSSRPQFCKWPCECPPSP--PRCPLGVSLITDGC-------ECCK------ 45
           | ||:||       | :     ||      |      : | |       ||
Db    4686 DCTPSPLCQLIKDSLFAQCHALVPPQHYYDAC------VFDSCFMPGSSLECASLQAYAA 4739

Qy      46 MCAQQL----GDNCTEAAI---CDPHR--------------------------GLYCDY 71
           :||||       | |  |    |  ||                         | :|
Db    4740 LCAQQNICLDWRNHTHGACLVECPSHREYQACGPAEEPTCKSSSSQQNNTVLVEGCFCPE 4799

Qy      72 SGDRPRYAIGVCAQVVGVGCV-LDGVRYNNGQSFQPNCKYNCTCIDGAVG--CTP----- 123
             || |    |   |||  | |    |: |: :|| || |::|  |  | |
Db    4800 G--TMNYAPGFDVCVKTCGCVGPDNVPREFGEHFEFDCK-NCVCLEGGSGIICQPKRCSQ 4856

Qy     124 ------------LCLRVRP-------------------PRLWCP---HPRRVSIPGHCC 148
                       |   | |                   | : ||    :   :|| ||
Db    4857 KPVTHCVEDGTYLATEVNPADTCCNITVCKCNTSLCKEKPSV-CPLGFEVKSKMVPGRCC 4915

Qy     149 EQWVCEDDAKRPRKTAPRDTGA-------FDAVGEVEAWHR---NCIAYTSPWSPCSTSC 198
             : ||          |    |:        | |   :  :    | || |    ||:|||
Db    4916 PFYWCESKGVCVHGNAEYQPGSPVYSSKCQDCVCTDKVDNNTLLNVIACTH--VPCNTSC 4973

Qy     199 GLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGK------------KCLAVYQ 246
             |    :     :|  : |   | ::  | : | ::| |              |: ::
Db    4974 SPGF--ELMEAPGECCKKCEQTHCIIKRPD-NQHVILKPGDFKSDPKNNCTFFSCVKIHN 5030

Qy     247 PEASM-------NFTLAGCI-STRSYQPKYC---------------------------- 269
             |         ||  : ||  : :: |  |
Db    5031 QLISSVSNITCPNFDASICIPGSITFMPNGCCKTCTPRNETRVPCSTVPVTTEVSYAGCT 5090

Qy     270 ---------GVC------------MDN--RCCIPYKSKTIDVSFQCPDGLGFSRQVLWIN 306
                    | |             :|:   ||   |:   :|   ||:|   :     |
Db    5091 KTVLMNHCSGSCGTFVMYSAKAQALDHSCSCCKEEKTSQREVVLSCPNGGSLTHTYTHIE 5150

Qy     307 ACFC 310
           :| |
Db    5151 SCQC 5154


RESULT  15
NTC1_MOUSE
ID   NTC1_MOUSE     STANDARD;      PRT;  2531 AA.
AC   Q01705;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1 PRECURSOR (MOTCH PROTEIN).
GN   NOTCH1 OR MOTCH.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=93194170; PubMed=8449489;
RA   Franco del Amo F., Gendron-Maguire M., Swiatek P.J., Jenkins N.A.,
RA   Copeland N.G., Gridley T.;
RT   "Cloning, analysis, and chromosomal localization of Notch-1, a mouse
RT   homolog of Drosophila Notch.";
RL   Genomics 15:259-264(1993).
RN   [2]
RP   SEQUENCE OF 1551-2170 FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=93048835; PubMed=1425352;
RA   Franco del Amo F., Smith D.E., Swiatek P.J., Gendron-Maguire M.,
RA   Greenspan R.J., McMahon A.P., Gridley T.;
RT   "Expression pattern of Motch, a mouse homolog of Drosophila Notch,
RT   suggests an important role in early postimplantation mouse
RT   development.";
RL   Development 115:737-744(1992).
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED ALMOST UNIFORMLY IN EARLY EMBRYOS.
CC   -!- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC   -!- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC   -!- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z11886; CAA77941.1; -.
DR   HSSP; P00740; 1IXA.
DR   MGD; MGI:97363; Notch1.
DR   InterPro; IPR000152; -.
DR   InterPro; IPR000561; -.
DR   InterPro; IPR000800; -.
DR   InterPro; IPR001438; -.
DR   InterPro; IPR001881; -.
DR   InterPro; IPR002110; -.
DR   Pfam; PF00008; EGF; 35.
DR   Pfam; PF00023; ank; 6.
DR   Pfam; PF00066; notch; 3.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PROSITE; PS50088; ANK_REPEAT; 2.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 22.
DR   PROSITE; PS00022; EGF_1; 34.
DR   PROSITE; PS01186; EGF_2; 27.
DR   PROSITE; PS01187; EGF_CA; 21.
KW   Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;
KW   Transmembrane; Signal; Glycoprotein.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19   2531       NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1.
FT   DOMAIN       19   1725       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1726   1746       POTENTIAL.
FT   DOMAIN     1747   2531       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       20     58       EGF-LIKE 1.
FT   DOMAIN       59     99       EGF-LIKE 2.
FT   DOMAIN      102    139       EGF-LIKE 3.
FT   DOMAIN      140    176       EGF-LIKE 4.
FT   DOMAIN      178    216       EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      218    255       EGF-LIKE 6.
FT   DOMAIN      257    293       EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      295    333       EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      335    371       EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      372    410       EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      412    450       EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      452    488       EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      490    526       EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      528    564       EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      566    601       EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      603    639       EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      641    676       EGF-LIKE 17, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      678    714       EGF-LIKE 18, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      716    751       EGF-LIKE 19, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      753    789       EGF-LIKE 20, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      791    827       EGF-LIKE 21, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      829    867       EGF-LIKE 22.
FT   DOMAIN      869    905       EGF-LIKE 23, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      907    943       EGF-LIKE 24.
```

```
FT   DOMAIN      945    981       EGF-LIKE 25, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      983   1019       EGF-LIKE 26.
FT   DOMAIN     1021   1057       EGF-LIKE 27, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     1059   1095       EGF-LIKE 28.
FT   DOMAIN     1097   1143       EGF-LIKE 29.
FT   DOMAIN     1145   1181       EGF-LIKE 30, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     1183   1219       EGF-LIKE 31, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     1221   1265       EGF-LIKE 32, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     1267   1305       EGF-LIKE 33.
FT   DOMAIN     1307   1346       EGF-LIKE 34.
FT   DOMAIN     1348   1384       EGF-LIKE 35.
FT   DOMAIN     1387   1426       EGF-LIKE 36.
FT   DOMAIN     1449   1462       CYS-RICH.
FT   REPEAT     1445   1480       LIN/NOTCH 1.
FT   REPEAT     1481   1522       LIN/NOTCH 2.
FT   REPEAT     1523   1562       LIN/NOTCH 3.
FT   REPEAT     1917   1947       ANK 1.
FT   REPEAT     1949   1979       ANK 2.
FT   REPEAT     1983   2012       ANK 3.
FT   REPEAT     2016   2045       ANK 4.
FT   REPEAT     2049   2078       ANK 5.
FT   DISULFID     24     37       BY SIMILARITY.
FT   DISULFID     31     46       BY SIMILARITY.
FT   DISULFID     63     74       BY SIMILARITY.
FT   DISULFID     68     87       BY SIMILARITY.
FT   DISULFID     89     98       BY SIMILARITY.
FT   DISULFID    106    117       BY SIMILARITY.
FT   DISULFID    111    127       BY SIMILARITY.
FT   DISULFID    129    138       BY SIMILARITY.
FT   DISULFID    144    155       BY SIMILARITY.
FT   DISULFID    149    164       BY SIMILARITY.
FT   DISULFID    166    175       BY SIMILARITY.
FT   DISULFID    182    195       BY SIMILARITY.
FT   DISULFID    189    204       BY SIMILARITY.
FT   DISULFID    206    215       BY SIMILARITY.
FT   DISULFID    222    233       BY SIMILARITY.
FT   DISULFID    227    243       BY SIMILARITY.
FT   DISULFID    245    254       BY SIMILARITY.
FT   DISULFID    261    272       BY SIMILARITY.
FT   DISULFID    266    281       BY SIMILARITY.
FT   DISULFID    283    292       BY SIMILARITY.
FT   DISULFID    299    312       BY SIMILARITY.
FT   DISULFID    306    321       BY SIMILARITY.
FT   DISULFID    323    332       BY SIMILARITY.
FT   DISULFID    339    350       BY SIMILARITY.
FT   DISULFID    344    359       BY SIMILARITY.
FT   DISULFID    361    370       BY SIMILARITY.
FT   DISULFID    376    387       BY SIMILARITY.
FT   DISULFID    381    398       BY SIMILARITY.
FT   DISULFID    400    409       BY SIMILARITY.
FT   DISULFID    416    429       BY SIMILARITY.
FT   DISULFID    423    438       BY SIMILARITY.
FT   DISULFID    440    449       BY SIMILARITY.
FT   DISULFID    456    467       BY SIMILARITY.
FT   DISULFID    461    476       BY SIMILARITY.
FT   DISULFID    478    487       BY SIMILARITY.
FT   DISULFID    494    505       BY SIMILARITY.
FT   DISULFID    499    514       BY SIMILARITY.
FT   DISULFID    516    525       BY SIMILARITY.
FT   DISULFID    532    543       BY SIMILARITY.
FT   DISULFID    537    552       BY SIMILARITY.
FT   DISULFID    554    563       BY SIMILARITY.
FT   DISULFID    570    580       BY SIMILARITY.
FT   DISULFID    575    589       BY SIMILARITY.
FT   DISULFID    591    600       BY SIMILARITY.
FT   DISULFID    607    618       BY SIMILARITY.
FT   DISULFID    612    627       BY SIMILARITY.
FT   DISULFID    629    638       BY SIMILARITY.
FT   DISULFID    645    655       BY SIMILARITY.
FT   DISULFID    650    664       BY SIMILARITY.
FT   DISULFID    666    675       BY SIMILARITY.
FT   DISULFID    682    693       BY SIMILARITY.
FT   DISULFID    687    702       BY SIMILARITY.
FT   DISULFID    704    713       BY SIMILARITY.
FT   DISULFID    720    730       BY SIMILARITY.
FT   DISULFID    725    739       BY SIMILARITY.
FT   DISULFID    741    750       BY SIMILARITY.
FT   DISULFID    757    768       BY SIMILARITY.
FT   DISULFID    762    777       BY SIMILARITY.
FT   DISULFID    779    788       BY SIMILARITY.
FT   DISULFID    795    806       BY SIMILARITY.
FT   DISULFID    800    815       BY SIMILARITY.
FT   DISULFID    817    826       BY SIMILARITY.
FT   DISULFID    833    844       BY SIMILARITY.
FT   DISULFID    838    855       BY SIMILARITY.
FT   DISULFID    857    866       BY SIMILARITY.
FT   DISULFID    873    884       BY SIMILARITY.
FT   DISULFID    878    893       BY SIMILARITY.
FT   DISULFID    895    904       BY SIMILARITY.
FT   DISULFID    911    922       BY SIMILARITY.
FT   DISULFID    916    931       BY SIMILARITY.
FT   DISULFID    933    942       BY SIMILARITY.
FT   DISULFID    987    998       BY SIMILARITY.
FT   DISULFID    992   1007       BY SIMILARITY.
FT   DISULFID   1009   1018       BY SIMILARITY.
FT   DISULFID   1025   1036       BY SIMILARITY.
FT   DISULFID   1030   1045       BY SIMILARITY.
FT   DISULFID   1047   1056       BY SIMILARITY.
FT   DISULFID   1063   1074       BY SIMILARITY.
FT   DISULFID   1068   1083       BY SIMILARITY.
FT   DISULFID   1085   1094       BY SIMILARITY.
FT   DISULFID   1101   1122       BY SIMILARITY.
FT   DISULFID   1116   1131       BY SIMILARITY.
FT   DISULFID   1133   1142       BY SIMILARITY.
FT   DISULFID   1149   1160       BY SIMILARITY.

  Query Match             7.7%;  Score 150;  DB 1;  Length 2531;
  Best Local Similarity   21.6%;  Pred. No. 8.9e-05;
  Matches   90;  Conservative   31;  Mismatches  135;  Indels  160;  Gaps   26;

Qy      21 CECPPS--PPRCPLGVS-------LITDGCEC--------CKMCAQQLGDNCTEAAI-CD 62
           | |||    | |   |:       |   | |         |  ||   | :|    : |
Db     164 CRCPPGFHGPTCRQDVNECSQNPGLCRHGGHCHNEIGSYRCACCATHTGPHCELPYVPCS 223

Qy      63 P---HRGLYCDYSGDR-------PRYAIGVCAQVV----------GVGCVLDGVR-YN-- 99
           |     |  | :||         | :|   | : |          |  || |||  ||
Db     224 PSPCQNGATCRPTGDTTHECACLPGFAGQNCEENVDDCPGNNCKNGGACV-DGVNTYNCR 282

Qy     100 -----NGQ---------SFQPN-CK----------YNCTCIDGAVG----------CTP 123
               ||             || |:          ||| |::|  |           :
Db     283 CPPEVTGQYCTEDVDECQLMPNACQNAGTCHNTHGGYNCVCVNGWTGEDCSENIDDCASA 342

Qy     124 LCL-------RVRPPRLWCPHPR--------RVSIPGHCCEQWVCEDDAKRPRKTAPRDT 168
           |         ||      ||| |          |   | |   |: :    ::     :
Db     343 ACFQGATCHDRVASFYCECPHGRTGLLCHLKHACISNPCNEGSNCDTNPVNGKRICTCPS 402

Qy     169 GAFDAVGEVEAWHRNCIAYTSPWSPCS---TSCGLGVS---------TRISNVNAQCWP- 215
           |              || |   ||     | || :           : :   ||
Db     403 G----------------YTGP--ACSQDVDECDLGANRCEHAGKCLNTLGSFECQCLQG 443

Qy     216 ------EQESRLCNLRPCDVDIHTLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYC 269
                | :   |   ||  |   | : |:     :|           ||    |: ||
Db     444 YTGPGCEIDVNECISNPCQNDATCLDQIGE-----FQ----------CICMPGYEGVYC 487

Qy     270 GV---------CMDNRCCIPYKSKTIDVSFQCPDGLGFSRQVLWINACFCNLSCRN 316
           :          |: |  |:     |  :   ||| |          :: |  :  |:|
Db     488 EINTDECASSPCLHNGHCM---DKIHEFQCQCPKGFNGHLCQYDVDEC-ASTPCKN 539


Search completed: July 26, 2001, 08:38:01
Job time: 113 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 26, 2001, 08:34:33 ; Search time 24.7 Seconds
(without alignments)
1794.420 Million cell updates/sec

Title: US-09-325-019-2
Perfect score: 1937
Sequence: 1 DFTPAPLEDTSSRPQFCKWP.....NPNDIFADLESYPDFSEIAN 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_16:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1937	100.0	367	4	O95388	O95388 homo sapien
2	1675	86.5	367	11	O54775	O54775 mus musculu
3	1364.5	70.4	280	4	Q9HCS3	Q9hcs3 homo sapien
4	860.5	44.4	349	6	Q9GL71	Q9gl71 bos taurus
5	852.5	44.0	347	13	Q9PT80	Q9pt80 notophthalm
6	844	43.6	354	4	O95389	O95389 homo sapien
7	836.5	43.2	347	11	Q9R1E9	Q9r1e9 rattus norv
8	835.5	43.1	347	11	Q9WVS1	Q9wvs1 rattus norv
9	830	42.8	343	13	O42607	O42607 xenopus lae
10	827.5	42.7	349	6	O97765	O97765 sus scrofa
11	815.5	42.1	331	4	O95958	O95958 homo sapien
12	780.5	40.3	379	11	Q9ES72	Q9es72 rattus norv
13	779.5	40.2	379	11	Q9WTM9	Q9wtm9 rattus norv
14	762.5	39.4	381	4	O43775	O43775 homo sapien
15	719.5	37.1	351	11	Q9QZQ5	Q9qzq5 rattus norv
16	555.5	28.7	334	4	Q9UID7	Q9uid7 homo sapien
17	498	25.7	251	11	Q9Z0G4	Q9z0g4 mus musculu
18	490.5	25.3	250	4	O76076	O76076 homo sapien
19	477.5	24.7	250	11	Q9JHC6	Q9jhc6 rattus norv
20	462.5	23.9	176	13	Q9PSS6	Q9pss6 gallus gall
21	328.5	17.0	128	11	Q9R2C0	Q9r2c0 rattus norv
22	305	15.7	100	4	Q9UDL6	Q9udl6 homo sapien
23	275	14.2	113	11	Q9Z164	Q9z164 rattus norv
24	218	11.3	74	11	Q9WUW4	Q9wuw4 rattus norv
25	182.5	9.4	77	4	Q9UDE4	Q9ude4 homo sapien
26	174.5	9.0	179	5	Q9VVK0	Q9vvk0 drosophila
27	174.5	9.0	2327	13	Q9IBG7	Q9ibg7 xenopus lae
28	169.5	8.8	4123	4	O75851	O75851 homo sapien
29	161.5	8.3	1637	6	Q9XSV8	Q9xsv8 bos taurus
30	161	8.3	925	4	Q9H318	Q9h318 homo sapien
31	161	8.3	1036	4	Q9NZV1	Q9nzv1 homo sapien
32	160.5	8.3	430	5	Q26424	Q26424 crassostrea
33	158	8.2	70	13	Q9DF21	Q9df21 scyliorhinu
34	153.5	7.9	3680	5	Q9VR08	Q9vr08 drosophila
35	149	7.7	543	5	Q9VJU5	Q9vju5 drosophila
36	149	7.7	620	5	Q9NKD8	Q9nkd8 drosophila
37	148.5	7.7	685	6	Q9TTS5	Q9tts5 bos taurus
38	146	7.5	765	5	Q9NL50	Q9nl50 sarcophaga
39	146	7.5	988	6	O97867	O97867 sus scrofa
40	145.5	7.5	1042	4	Q13792	Q13792 homo sapien
41	145.5	7.5	1081	4	O76065	O76065 homo sapien
42	145	7.5	1028	11	Q9JLL0	Q9jll0 mus musculu
43	144	7.4	1444	5	Q17591	Q17591 caenorhabdi
44	142.5	7.4	2843	4	Q9Y6R7	Q9y6r7 homo sapien
45	142	7.3	1111	5	Q9XWD6	Q9xwd6 caenorhabdi

ALIGNMENTS

RESULT 1
O95388
ID O95388 PRELIMINARY; PRT; 367 AA.
AC O95388;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE CONNECTIVE TISSUE GROWTH FACTOR RELATED PROTEIN WISP-1.
GN WISP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ADULT LUNG, AND FETAL KIDNEY;
RX MEDLINE=99061933; PubMed=9843955;
RA Pennica D., Swanson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,
RA Brush J., Taneyhill L.A., Deuel B., Lew M., Watanabe C., Cohen R.L.,
RA Melham M.F., Finley G.G., Quirke P., Goddard A.D., Hillan K.J.,
RA Gurney A.L., Botstein D., Levine A.J.;
RT "WISP genes are members of the connective tissue growth factor family
RT that are up-regulated in wnt-1-transformed cells and aberrantly
RT expressed in human colon tumors.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Blechschmidt K., Kalaydjieva L., Goodman R., Gresham D., Baas F.,
RA Jonge R.d., Schilhabel M., Schattevoy R., Dette M., Menzel U.,
RA Rosenthal A.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF100779; AAC96321.1; -.
DR EMBL; AF192304; AAF22341.1; -.
DR InterPro; IPR000359; -.
DR InterPro; IPR000867; -.
DR InterPro; IPR000884; -.
DR InterPro; IPR001007; -.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; vwc; 1.
DR Pfam; PF00219; IGFBP; 1.
DR PROSITE; PS01185; CTCK_1; 1.

```
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS01208; VWFC; UNKNOWN_1.
DR   SMART; SM00209; TSP1; 1.
SQ   SEQUENCE   367 AA;  40331 MW;  9F29CA94D69C0502 CRC64;


  Query Match            100.0%;  Score 1937;  DB 4;  Length 367;
  Best Local Similarity  100.0%;  Pred. No. 9.2e-185;
  Matches  335;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     33 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 92

Qy     61 CDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     93 CDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVG 152

Qy    121 CTPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAW 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    153 CTPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAW 212

Qy    181 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    213 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 272

Qy    241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    273 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 332

Qy    301 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
          |||||||||||||||||||||||||||||||||||
Db    333 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 367


RESULT   2
O54775
ID   O54775      PRELIMINARY;      PRT;   367 AA.
AC   O54775;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   ELM1.
GN   ELM1 OR WISP1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HEN;
RX   MEDLINE=98119879; PubMed=9449709;
RA   Hashimoto Y., Shindo-okada N., Tani M., Nagamachi Y., Takeuchi K.,
RA   Shiroishi T., Toma H., Yokota J.;
RT   "Expression of the Elm1 gene, a novel gene of the CCN (connective
RT   tissue growth factor, Cyr61/Cef10, and neuroblastoma overexpressed
RT   gene) family, suppresses In vivo tumor growth and metastasis of K-1735
RT   murine melanoma cells.";
RL   J. Exp. Med. 187:289-296(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=MAMMARY;
RX   MEDLINE=99061933; PubMed=9843955;
RA   Pennica D., Swanson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,
RA   Brush J., Taneyhill L.A., Deuel B., Lew M., Watanabe C., Cohen R.L.,
RA   Melham M.F., Finley G.G., Quirke P., Goddard A.D., Hillan K.J.,
RA   Gurney A.L., Botstein D., Levine A.J.;
RT   "WISP genes are members of the connective tissue growth factor family
RT   that are up-regulated in wnt-1-transformed cells and aberrantly
RT   expressed in human colon tumors.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).
DR   EMBL; AB004873; BAA24949.1; -.
DR   EMBL; AF100777; AAC96319.1; -.
DR   MGD; MGI:1197008; Wisp1.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001007; -.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00093; vwc; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS01208; VWFC; UNKNOWN_1.
DR   SMART; SM00041; CT; 1.
SQ   SEQUENCE   367 AA;  40702 MW;  3B7C0569EFAB5E96 CRC64;


  Query Match            86.5%;  Score 1675;  DB 11;  Length 367;
  Best Local Similarity  85.3%;  Pred. No. 1.1e-158;
  Matches  285;  Conservative   20;  Mismatches   29;  Indels    0;  Gaps    0;

Qy      2 FTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAIC 61
          |||||||:|::||:|||||||||| ||||||||||||||||||:|||||||||||||||
Db     34 FTPAPLEETTTRPEFCKWPCECPQSPPRCPLGVSLITDGCECCKICAQQLGDNCTEAAIC 93

Qy     62 DPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGC 121
          |||||||||||||||||||||||||||||||||||| ||:|||||||:|||||| |||
Db     94 DPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYTNGESFQPNCRYNCTCIDGTVGC 153

Qy    122 TPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWH 181
          |||||  |||||||  || | :|| |||||||:|||:|||:||  || || | | ||  :
Db    154 TPLCLSPRPPRLWCRQPRHVRVPGQCCEQWVCDDDARRPRQTALLDTRAFAASGAVEQRY 213

Qy    182 RNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKKC 241
          |||||||||||||:||||:|||||||||:||||||||||||||||||   |||||||
Db    214 ENCIAYTSPWSPCSTTCGLGISTRISNVNARCWPEQESRLCNLRPCDVDIQLHIKAGKKC 273

Qy    242 LAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSRQ 301
          ||||||| : |||||||:|||:|:|||||||| ||||||||||||| | ||||:| |||||
Db    274 LAVYQPEEATNFTLAGCVSTRTYRPKYCGVCTDNRCCIPYKSKTISVDFQCPEGPGFSRQ 333

Qy    302 VLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
          |||||||||||||||||||||||||||| ||||
Db    334 VLWINACFCNLSCRNPNDIFADLESYPDFEEIAN 367


RESULT   3
Q9HCS3
ID   Q9HCS3      PRELIMINARY;      PRT;   280 AA.
AC   Q9HCS3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   WISP-1 VARIANT.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tanaka S., Sugimachi K.;
RT   "Human WISP-1 variant.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB034725; BAB17849.1; -.
SQ   SEQUENCE   280 AA;  30697 MW;  26B254D4A060738E CRC64;


  Query Match            70.4%;  Score 1364.5;  DB 4;  Length 280;
  Best Local Similarity  73.7%;  Pred. No. 6.4e-128;
  Matches  247;  Conservative    0;  Mismatches    1;  Indels   87;  Gaps    1;

Qy      1 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     33 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 92
```

```
Qy      61 CDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVG 120
           ||||||||||||||||||||||||
Db      93 CDPHRGLYCDYSGDRPRYAIGVCAH---------------------------------- 117

Qy     121 CTPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAW 180
                                                              ||||||||
Db     118 ----------------------------------------------------AVGEVEAW 125

Qy     181 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     126 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 185

Qy     241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     186 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 245

Qy     301 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
           |||||||||||||||||||||||||||||||||||
Db     246 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 280
```

```
RESULT   4
Q9GL71
ID   Q9GL71      PRELIMINARY;      PRT;   349 AA.
AC   Q9GL71;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR.
GN   CTGF.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Mathahs M., Schwitters C., Hove M., Rupp S., Erondu N.E.;
RT   "Bovine connective tissue growth factor, organization of the
RT   chromosomal gene and demonstration of promoter activity.";
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF309555; AAG30290.1; -.
SQ   SEQUENCE   349 AA;  37924 MW;  5FFC8EE83EFB4F99 CRC64;
```

```
  Query Match           44.4%;  Score 860.5;  DB 6;  Length 349;
  Best Local Similarity   45.7%;  Pred. No. 1.1e-77;
  Matches  155;  Conservative   52;  Mismatches  103;  Indels   29;  Gaps    7;

Qy      12 SRP---QFCKWPCECPPSP-PRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
           |||   | |  ||:||  | |||| ||||: ||| ||::||:|| : |||   ||||:||
Db      21 SRPASSQDCSAPCQCPAGPAPRCPAGVSLVLDGCGCCRVCAKQLSELCTERDPCDPHKGL 80

Qy      68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
           :||: |      ||||    |  ||  |  | :|:||| :||| |||:||:||| ||| :
Db      81 FCDF-GSPANRKIGVCTAKDGAPCVFGGTVYQSGESFQSSCKYQCTCLDGSVGCVPLCSV 139

Qy     127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDD----------AKRPRKTAPRDTGAFDAVG 175
            || |   || |||| :|| |||:|||::           | ||  |   |    |
Db     140 DVRLPSPDCPFPRRVKLPGKCCEEWVCDEPKEHTVVGPALAAYRPEDTFGPDPTMIRA-- 197

Qy     176 EVEAWHRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLI 235
                ||:  |: || || :|||:|:|||::| || |  |::|||| :|||: |:   |
Db     198 -------NCLVQTTEWSACSKTCGMGISTRVTNDNAFCRLEKQSRLCMVRPCEADLEENI 250

Qy     236 KAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDG 295
           | ||||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:||||
Db     251 KKGKKCIRTPKISKPIKFELSGCTSMKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKCPDG 310

Qy     296 LGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
               : :::|  | |: :|   ||||  |   : | | :
Db     311 EVMKKSMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 349
```

```
RESULT   5
Q9PT80
ID   Q9PT80      PRELIMINARY;      PRT;   347 AA.
AC   Q9PT80;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CONNECTIVE TISSUE GROWTH FACTOR.
GN   CTGF.
OS   Notophthalmus viridescens (Eastern newt) (Triturus viridescens).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae;
OC   Notophthalmus.
OX   NCBI_TaxID=8316;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=FORELIMB BLASTEMA;
RA   Gates P.B.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=FORELIMB BLASTEMA;
RX   MEDLINE=99033008; PubMed=9813273;
RA   Cash D.E., Gates P.B., Imokawa Y., Brockes J.P.;
RT   "Identification of newt connective tissue growth factor as a target of
RT   retinoid regulation in limb blastenal cells.";
RL   Gene 222:119-124(1998).
DR   EMBL; AJ271167; CAB65965.1; -.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001007; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00093; vwc; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS00222; IGF_BINDING; 1.
DR   PROSITE; PS01208; VWFC; UNKNOWN_1.
DR   SMART; SM00041; CT; 1.
SQ   SEQUENCE   347 AA;  38098 MW;  3B7E2399F27672C1 CRC64;
```

```
  Query Match           44.0%;  Score 852.5;  DB 13;  Length 347;
  Best Local Similarity   44.5%;  Pred. No. 6.8e-77;
  Matches  149;  Conservative   53;  Mismatches  102;  Indels   31;  Gaps    5;

Qy      15 QFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGLYCDYSGD 74
           | |   | ||  || || | ||: ||| |||:||:|||: |||  :|||||||:||: |
Db      26 QDCSGECRCPNKPPECPAGTSLVMDGCGCCKVCAKQLGELCTEKDVCDPHRGLFCDF-GS 84

Qy      75 RPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-LRVRPPRL 133
           |    ||||    |  ||  |: | :|:||| :||| |||:|| ||| ||| : || |
Db      85 RVNKKIGVCTAKDGAPCVFGGMVYRSGESFQSSCKYQCTCLDGGVGCVPLCGVDVRLPSP 144

Qy     134 WCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---------- 182
            || ||||   | |||:|||:   :: |            ||   | :|
Db     145 DCPFPRRVKQLGKCCEEWVCDQPKEQTR-----------VGPALAVYRQEETYGPDSSL 192

Qy     183 ---NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGK 239
              ||:  |: || || :|||:|:|||::| |  |  |::|||| :|||: |:   || ||
Db     193 MRANCLVQTTEWSACSKTCGMGISTRVTNDNEMCRLEKQSRLCMVRPCEADLEENIKKGK 252

Qy     240 KCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFS 299
           ||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:||||
Db     253 KCIRTPKISKPVKFELSGCTSVKTYRAKFCGVCTDGRCCTPHRTATLPVEFKCPDGEVMK 312

Qy     300 RQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
           :::::|  | |: :|   ||||  :   : | | :
Db     313 KKMMFIKTCACHYNCPGDNDIFESMYYRKMYGDMA 347
```

```
RESULT   6
O95389
ID   O95389      PRELIMINARY;      PRT;   354 AA.
AC   O95389;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CONNECTIVE TISSUE GROWTH FACTOR RELATED PROTEIN WISP-3.
GN   WISP3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BONE MARROW, AND FETAL KIDNEY;
RX   MEDLINE=99061933; PubMed=9843955;
RA   Pennica D., Swanson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,
RA   Brush J., Taneyhill L.A., Deuel B., Lew M., Watanabe C., Cohen R.L.,
RA   Melham M.F., Finley G.G., Quirke P., Goddard A.D., Hillan K.J.,
RA   Gurney A.L., Botstein D., Levine A.J.;
RT   "WISP genes are members of the connective tissue growth factor family
RT   that are up-regulated in wnt-1-transformed cells and aberrantly
RT   expressed in human colon tumors.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).
DR   EMBL; AF100781; AAC96323.1; -.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS00222; IGF_BINDING; 1.
DR   SMART; SM00041; CT; 1.
SQ   SEQUENCE   354 AA;  39292 MW;  67F48D0D5C2F5EE3 CRC64;

  Query Match             43.6%;  Score 844;  DB 4;  Length 354;
  Best Local Similarity   43.2%;  Pred. No. 4.9e-76;
  Matches  146;  Conservative   54;  Mismatches  102;  Indels   36;  Gaps    5;

Qy       4 PAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDP 63
           |  : |    | ||| |||:||   |||| ||||: ||| |||:||:| |: | || :|||
Db      35 PGEVSDAPQRKQFCHWPCKCPQQKPRCPPGVSLVRDGCGCCKICAKQPGEICNEADLCDP 94

Qy      64 HRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTP 123
           |:|||||| |||||  |||| :| |||  : | |:||| ||||  ::| |: ||:||||
Db      95 HKGLYCDYSVDRPRYETGVCAYLVAVGCEFNQVHYHNGQVFQPNPLFSCLCVSGAIGCTP 154

Qy     124 LCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEA---- 179
           | :    |:|              ||      ||  :|:   :       : ::
Db     155 LFI----PKL---------AGSHC-------SGAKGGKKSDQSNCSLEPLLQQLSTSYKT 194

Qy     180 ----------WHRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDV 229
                     | : |:   : |:||| :||:|:| |::| |: |   :| ||| ::|||
Db     195 MPAYRNLPLIWKKKCLVQATKWTPCSRTCGMGISNRVTNENSNCEMRKEKRLCYIQPCDS 254

Qy     230 DIHTLIK--AGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTID 287
           :|   ||   || |   :|   :  |  :|| ||:||:| :||:|:| ||||| ||| |
Db     255 NILKTIKIPKGKTCQPTFQLSKAEKFVFSGCSSTQSYKPTFCGICLDKRCCIPNKSKMIT 314

Qy     288 VSFQCPDGLGFSRQVLWINACFCNLSCRNPNDIFADLE 325
           : | ||:   |  ::||| :| |  :|| | |||::|:
Db     315 IQFDCPNEGSFKWKMLWITSCVCQRNCREPGDIFSELK 352


RESULT   7
Q9R1E9
ID   Q9R1E9      PRELIMINARY;      PRT;   347 AA.
AC   Q9R1E9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CONNECTIVE TISSUE GROWTH FACTOR.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Xu J., Smock S.L., Rosenzweig A.B., Odgren P.R., Safadi F.F.,
RA   Marks S.C. Jr., Owen T.A., Popoff S.N.;
RT   "Cloning of the cDNA for Rat Connective Tissue Growth Factor (CTGF):
RT   Implications for Skeletal Development.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF120275; AAD39132.1; -.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001007; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00093; vwc; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS00222; IGF_BINDING; 1.
DR   PROSITE; PS01208; VWFC; 1.
DR   SMART; SM00041; CT; 1.
SQ   SEQUENCE   347 AA;  37756 MW;  CFBE1A19766B7B16 CRC64;

  Query Match             43.2%;  Score 836.5;  DB 11;  Length 347;
  Best Local Similarity   44.4%;  Pred. No. 2.7e-75;
  Matches  152;  Conservative   55;  Mismatches  100;  Indels   35;  Gaps    8;

Qy      12 SRP---QFCKWPCEC-PPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
           :||   | |   |:|   : |||| |||| ||| ||::||:||| ||| |||| :||
Db      19 TRPATGQDCSAQCQCAAEAAPRCPAGVSLVLDGCGCCRVCAKQLGELCTERDPCDPHKGL 78

Qy      68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
           :||: |      ||||    |  ||  |  | :|:||| :||| |||:|||||| ||| :
Db      79 FCDF-GSPANRKIGVCTAKDGAPCVFGGSVYRSGESFQSSCKYQCTCLDGAVGCVPLCSM 137

Qy     127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
           || |   || |||| :|| |||:|||::          |:|      ||   | :|
Db     138 DVRLPSPDCPFPRRVKLPGKCCEEWVCDE---------PKDR---TVVGPALAAYRLEDT 185

Qy     183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                     ||:  |: || || :||:|:|||::| |  |  |::|||| :|||: |:
Db     186 FGPDPTMMRANCLVQTTEWSACSKTCGMGISTRVTNDNTFCRLEKQSRLCMVRPCEADLE 245

Qy     233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
           || ||||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db     246 ENIKKGKKCIRTPKIAKPVKFELSGCTSVKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKC 305

Qy     293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
           |||    : :::|  | |: :|   ||||  |   : | | :
Db     306 PDGEIMKKNMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 347


RESULT   8
Q9WVS1
ID   Q9WVS1      PRELIMINARY;      PRT;   347 AA.
AC   Q9WVS1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CONNECTIVE TISSUE GROWTH FACTOR.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tezuka K., Tamatani T.;
RT   "Rattus norvegicus connective tissue growth factor.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB023068; BAA82125.1; -.
DR   InterPro; IPR000072; -.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001007; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00093; vwc; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   ProDom; PD001629; -; 1.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS00222; IGF_BINDING; 1.
DR   PROSITE; PS01208; VWFC; 1.
DR   SMART; SM00041; CT; 1.
SQ   SEQUENCE   347 AA;  37837 MW;  6A69511DE72FBF1C CRC64;


  Query Match             43.1%;  Score 835.5;  DB 11;  Length 347;
  Best Local Similarity   44.4%;  Pred. No. 3.4e-75;
  Matches  152;  Conservative   55;  Mismatches  100;  Indels   35;  Gaps    8;

Qy      12 SRP---QFCKWPCECP-PSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
           :||   | |   |:|   : |||| |||| ||| ||::||:|||: |||   ||||:||
Db      19 TRPATGQDCSAQCQCAREAAPRCPAGVSLVLDGCGCCRVCAKQLGELCTERDPCDPHKGL 78

Qy      68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
           :||: |      ||||    |  ||  |  | :|:||| :||| |||:|||||| ||| :
Db      79 FCDF-GSPANRKIGVCPAKDGAPCVFGGSVYRSGESFQSSCKYQCTCLDGAVGCVPLCSM 137

Qy     127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
            || |   || |||| :|| |||:|||::          |:|      ||   | :|
Db     138 DVRLPSPDCPFPRRVKLPGKCCEEWVCDE---------PKDR---TVVGPALAAYRLEDT 185

Qy     183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                     ||:  |: || || :||:|:|||::| |  |  |::|||| :|||: |:
Db     186 FGPDPTMMRANCLVQTTEWSACSKTCGMGISTRVTNDNTFCRLEKQSRLCMVRPCEADLE 245

Qy     233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
           || ||||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db     246 ENIKKGKKCIRTPKIAKPVKFELSGCTSVKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKC 305

Qy     293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
           |||    : :::|  | |: :|   ||||  |   : | | :
Db     306 PDGEIMKKNMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 347


RESULT   9
O42607
ID   O42607      PRELIMINARY;      PRT;   343 AA.
AC   O42607;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CONNECTIVE TISSUE GROWTH FACTOR XCTGF.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ying Z., King M.L.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U43524; AAB67639.1; -.
DR   EMBL; U43523; AAB67638.1; -.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001007; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00093; vwc; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS01208; VWFC; 1.
DR   SMART; SM00041; CT; 1.
SQ   SEQUENCE   343 AA;  37966 MW;  93F221C5DB565A81 CRC64;


  Query Match             42.8%;  Score 830;  DB 13;  Length 343;
  Best Local Similarity   45.1%;  Pred. No. 1.2e-74;
  Matches  142;  Conservative   53;  Mismatches  112;  Indels    8;  Gaps    4;

Qy      12 SRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGLYCDY 71
           |  | |   |:||   | |  || :: ||| |||:|::|||: |||  :||||:||:||:
Db      19 SDAQECNGECQCPNKVPVCDPGVRMVQDGCGCCKVCSKQLGELCTERDVCDPHKGLFCDF 78

Qy      72 SGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-LRVRP 130
           | |    ||||    |  ||  |  | :|:||| :||| |||||| ||| ||| : :|
Db      79 -GSRVNRKIGVCTAREGAPCVFGGTVYRSGESFQSSCKYQCTCIDGGVGCVPLCSMDIRL 137

Qy     131 PRLWCPHPRRVSIPGHCCEQWVCEDDAKR----PRKTAPRDTGAFDAVGEVEAWHRNCIA 186
           |   || |||| :|| |||:|||:   :|    |  | |    :   :      ||:
Db     138 PSPECPFPRRVKLPGKCCEEWVCDQPQERTLVGPALPAFRMEETYGP--DPSLIRANCLV 195

Qy     187 YTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKKCLAVYQ 246
           |: || || :||:|:|||::| |  |  |::|||| :|||: |:   || ||||:   :
Db     196 QTTEWSACSKTCGMGISTRVTNDNEHCRLEKQSRLCMVRPCEADLEENIKKGKKCIRTPK 255

Qy     247 PEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSRQVLWIN 306
           : |  :|| | ::|: |:|||| | ||| |::: |: | |:|||    : :::|
Db     256 ISKPVKFEFSGCTSVKTYRAKFCGVCTDGRCCTPHRTATLPVEFKCPDGEVMKKNMMFIK 315

Qy     307 ACFCNLSCRNPNDIF 321
           | |: :|   ||||
Db     316 TCACHFNCPGDNDIF 330


RESULT  10
O97765
ID   O97765      PRELIMINARY;   PRT;   349 AA.
AC   O97765;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CONNECTIVE TISSUE GROWTH FACTOR.
GN   CTGF.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=UTERUS;
RA   Harding P.A., Brigstock D.R.;
RT   "Cloning and sequencing of a porcine connective tissue growth factor
RT   (CTGF) cDNA.";
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U70060; AAD00174.1; -.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001007; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00093; vwc; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   PROSITE; PS01185; CTCK_1; 1.
```

```
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS00222; IGF_BINDING; 1.
DR   PROSITE; PS01208; VWFC; 1.
DR   SMART; SM00041; CT; 1.
SQ   SEQUENCE   349 AA;  37946 MW;  35AB4275AC1D4B3A CRC64;


  Query Match             42.7%;  Score 827.5;  DB 6;  Length 349;
  Best Local Similarity   44.2%;  Pred. No. 2.1e-74;
  Matches  151;  Conservative   53;  Mismatches  103;  Indels   35;  Gaps    8;

Qy      12 SRP---QFCKWPCECPPSPPR-CPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGL 67
           |||   | |   |:|     | || ||| : :|| ||::||: ||| ||| | ||||:||
Db      21 SRPASGQDCSGQCQCAAGKRRACPAGVSSLLEGCGCCRLCAKHLGDLCTERAPCDPHKGL 80

Qy      68 YCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-L 126
           :||: |      ||||       ||  |  | :|:||| :||| |||:|||||| ||| :
Db      81 FCDF-GSPANRKIGVCTAKDCAPCVFGGTVYRSGESFQSSCKYQCTCLDGAVGCVPLCSM 139

Qy     127 RVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHR---- 182
           || |   || |||| :|| |||:|||::         |:|      ||   | :|
Db     140 DVRLPSPDCPFPRRVKLPGKCCEEWVCDE---------PKD---HTVVGPALAAYRLEDT 187

Qy     183 ----------NCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIH 232
                    ||:  |: || || :||:|:|||::| || |  |::|||| :|||: |:
Db     188 FGPDPTMMRANCLVQTTEWSACSKTCGMGISTRVTNDNASCRLEKQSRLCMVRPCEADLE 247

Qy     233 TLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQC 292
           || ||||:   :    : | |:|| | ::|: |:|||| | ||| |::: |: | |:|
Db     248 ENIKKGKKCIRTPKISKPVKFELSGCTSVKTYRAKFCGVCTDGRCCTPHRTTTLPVEFKC 307

Qy     293 PDGLGFSRQVLWINACFCNLSCRNPNDIFADL---ESYPDFS 331
           |||    : :::|  | |: :|   ||||  |   : | | :
Db     308 PDGEVMKKSMMFIKTCACHYNCPGDNDIFESLYYRKMYGDMA 349


RESULT  11
O95958
ID   O95958         PRELIMINARY;   PRT;   331 AA.
AC   O95958;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   DJ142L7.3 (CONNECTIVE TISSUE GROWTH FACTOR (NOV, GIG) LIKE PROTEIN)
DE   (FRAGMENT).
GN   DJ142L7.3 OR LIBC.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tubby B.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=MAMMARY GLAND;
RA   van Golen K.L., Davies S., Wu Z.F., Wang Y.F., Bucana C.D., Root H.,
RA   Chandrasekharappa S., Strawderman M., Ethier S.P., Merajver S.D.;
RT   "A novel putative IGF-binding, tumor suppressor protein, LIBC, and
RT   RhoC GTPase, are determinants of the inflammatory breast cancer
RT   phenotype.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z99289; CAB16556.1; -.
DR   EMBL; AF143679; AAD31517.1; -.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS00222; IGF_BINDING; 1.
DR   SMART; SM00041; CT; 1.
FT   NON_TER     331    331
SQ   SEQUENCE   331 AA;  36909 MW;  D109C2FDCA1DF549 CRC64;


  Query Match             42.1%;  Score 815.5;  DB 4;  Length 331;
  Best Local Similarity   43.2%;  Pred. No. 3.1e-73;
  Matches  143;  Conservative   53;  Mismatches   98;  Indels   37;  Gaps    6;

Qy      11 SSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGLYCD 70
           : | ||| |||:|| |||| ||||: ||| |||:||:| |: | || :||||:|||||
Db      20 AGRKQFCHWPCKCPQQKPRCPPGVSLVRDGCGCCKICAKQPGEICNEADLCDPHKGLYCD 79

Qy      71 YSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLCLRVRP 130
           || ||||| |||| |||  : | |:||| |||| ::| |: ||:||||| :
Db      80 YSVDRPRYETGVCA-CKSVGCEFNQVHYHNGQVFQPNPLFSCLCVSGAIGCTPLFI---- 134

Qy     131 PRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEA---------- 179
           |:|            ||         ||  :|:   :      : ::
Db     135 PKL---------AGSHC-------SGAKGGKKSDQSNCSLEPLLQQLSTSYKTMPAYRNL 178

Qy     180 ---WHRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIK 236
               | : |:   : |:||| :|||:|:| |:::| |: |   :| ||| ::||| :|   ||
Db     179 PLIWKKKCLVQATKWTPCSRTCGMGISNRVTNENSNCEMRKEKRLCYIQPCDSNILKTIK 238

Qy     237 --AGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPD 294
             || |   :|   :  |  :|| ||:|||:| :||:|:| ||||| ||| | : | ||:
Db     239 IPKGKTCQPTFQLSKAEKFVFSGCSSTQSYKPTFCGICLDKRCCIPNKSKMITIQFDCPN 298

Qy     295 GLGFSRQVLWINACFCNLSCRNPNDIFADLE 325
           |  ::||| :| |  :|| | |||::|:
Db     299 EGSFKWKMLWITSCVCQRNCREPGDIFSELK 329


RESULT  12
Q9ES72
ID   Q9ES72      PRELIMINARY;      PRT;   379 AA.
AC   Q9ES72;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CYSTEINE-RICH PROTEIN 61.
GN   CYR61.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG, AND PLACENTA;
RX   MEDLINE=20435857; PubMed=10852911;
RA   Albrecht C., von Der Kammer H., Mayhaus M., Klaudiny J., Schweizer M.,
RA   Nitsch R.M.;
RT   "Muscarinic acetylcholine receptors induce the expression of the
RT   immediate early growth regulatory gene CYR61.";
RL   J. Biol. Chem. 275:28929-28936(2000).
DR   EMBL; AF218568; AAG14964.1; -.
SQ   SEQUENCE   379 AA;  41687 MW;  62BF0BBA4C5AFDE9 CRC64;


  Query Match             40.3%;  Score 780.5;  DB 11;  Length 379;
  Best Local Similarity   39.9%;  Pred. No. 1.1e-69;
  Matches  142;  Conservative   57;  Mismatches  112;  Indels   45;  Gaps    6;

Qy      17 CKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGLYCDYSGDRP 76
           |   | ||   |:|  || |: ||| |||:||:|| ::|::   ||  :|| |:: |
Db      26 CPAACHCPLEAPKCAPGVGLVRDGCGCCKVCAKQLNEDCSKTQPCDHTKGLECNF-GASS 84

Qy      77 RYAIGVC-AQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-LRVRPPRLW 134
           |:| ||  |  |  :   | ||:|||||||: ||||||||| |||   :  | |
Db      85 TALKGICRAQSEGRPCEYNSRIYQNGESFQPNCKHQCTCIDGAVGCIPLCPQELSLPNLG 144

Qy     135 CPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHRN---------- 183
```

```
          ||:|| | : | |||:|||::|: :       |   |||  |||  .  |
Db    145 CPNPRLVKVSGQCCEEWVCDEDSIKDSLDDQDDLLGFDA-SEVELTRNNELIAIGKGSSL 203

Qy    184 ----------------------CIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESR 220
                                 ||  |: || || ||| |:|||::| |::|   :|:|
Db    204 KRLPVFGTEPRVLYNPLHAHGQKCIVQTTSWSQCSKSCGTGISTRVTNDNSECRLVKETR 263

Qy    221 LCNLRPCDVDIHTLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIP 280
          :| :|||   ::: :| ||||    :    : || ||| | : |:|||| |:| ||| |
Db    264 ICEVRPCGQPVYSSLKKGKKCSKTKKSPEPVRFTYAGCSSVKKYRPKYCGSCVDGRCCTP 323

Qy    281 YKSKTIDVSFQCPDGLGFSRQVLWINACFCNLSCRNPND-------IFADLESYPD 329
           :::|: : |:| ||  ||: |: | :| || :| :||:       :| |:  : |
Db    324 LQTRTVKMRFRCEDGEMFSKNVMMIQSCKCNYNCPHPNEASFRLYSLFNDIHKFRD 379


RESULT  13
Q9WTM9
ID   Q9WTM9      PRELIMINARY;      PRT;   379 AA.
AC   Q9WTM9;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CYR61 PRECURSOR.
GN   CYR61.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IZM; TISSUE=AORTA;
RA   Unoki H., Yonekura H., Furukawa K., Yamamoto H.;
RT   "Rat Cyr61 mRNA.";
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB015877; BAA78339.1; -.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001007; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00093; vwc; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS00222; IGF_BINDING; 1.
DR   PROSITE; PS01208; VWFC; 1.
DR   SMART; SM00041; CT; 1.
KW   Signal.
FT   SIGNAL       1     24       POTENTIAL.
FT   CHAIN       25    379       CYR61.
SQ   SEQUENCE   379 AA;  41728 MW;  D2ABAFD77B84762B CRC64;


  Query Match             40.2%;  Score 779.5;  DB 11;  Length 379;
  Best Local Similarity   39.9%;  Pred. No. 1.4e-69;
  Matches  142;  Conservative   56;  Mismatches  113;  Indels   45;  Gaps    6;

Qy      17 CKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGLYCDYSGDRP 76
           |   | ||   |:|  || |: ||| |||:||:|| ::|::   ||  :|| |:: |
Db      26 CPASCHCPLEAPKCAPGVGLVRDGCGCCKVCAKQLNEDCSKTQPCDHTKGLECNF-GANS 84

Qy      77 RYAIGVC-AQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-LRVRPPRLW 134
             |:| ||  |  |  :   | ||:||||||: |||||||||| |||   :  | |
Db      85 TALKGICRAQSEGRPCEYNSRIYQNGESFQPNCKHQCTCIDGAVGCIPLCPQELSLPNLG 144

Qy     135 CPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAWHRN---------- 183
           ||:|| | : | |||:|||::|: :       |   |||  |||    |
Db     145 CPNPRLVKVSGQCCEEWVCDEDSIKDSLDDQDDLLGFDA-SEVELTRNNELIATGKGSSL 203

Qy     184 ----------------------CIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESR 220
                                 ||  |: || || ||| |:|||::| | :|   :|:|
Db    204 KRLPVFGTEPRVLYNPLHAHGQKCIVQTTSWSQCSKSCGTGISTRVTNDNPECRLVKETR 263

Qy    221 LCNLRPCDVDIHTLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIP 280
          :| :|||   ::: :| ||||    :    : || ||| | : |:|||| |:| ||| |
Db    264 ICEVRPCGQPVYSSLKKGKKCSKTKKSPEPVRFTYAGCSSVKKYRPKYCGSCVDGRCCTP 323

Qy    281 YKSKTIDVSFQCPDGLGFSRQVLWINACFCNLSCRNPND-------IFADLESYPD 329
           :::|: : |:| ||  ||: |: | :| || :| :||:       :| |:  : |
Db    324 LQTRTVKMRFRCEDGEMFSKNVMMIQSCKCNYNCPHPNEASFRLYSLFNDIHKFRD 379


RESULT  14
O43775
ID   O43775      PRELIMINARY;      PRT;   381 AA.
AC   O43775;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CYR61 PROTEIN.
GN   CYR61.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98197344; PubMed=9536281;
RA   Martinerie C., Viegas-Pequignot E., Nguyen V.C., Perbal B.;
RT   "Chromosomal mapping and expression of the human cyr61 gene in tumour
RT   cells from the nervous system.";
RL   Mol. Pathol. 50:310-316(1997).
DR   EMBL; Y11307; CAA72167.1; -.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001007; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00093; vwc; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS00222; IGF_BINDING; 1.
DR   PROSITE; PS01208; VWFC; 1.
DR   SMART; SM00041; CT; 1.
SQ   SEQUENCE   381 AA;  42025 MW;  9883CE4886C4B430 CRC64;


  Query Match             39.4%;  Score 762.5;  DB 4;  Length 381;
  Best Local Similarity   39.7%;  Pred. No. 6.7e-68;
  Matches  142;  Conservative   56;  Mismatches  113;  Indels   47;  Gaps    8;

Qy      17 CKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGLYCDYSGDRP 76
           |   | ||   |:|  || |: ||| |||:||:|| ::|::   ||  :|| |:: |
Db      26 CPAACHCPLEAPKCAPGVGLVRDGCGCCKVCAKQLNEDCSKTQPCDHTKGLECNF-GASS 84

Qy      77 RYAIGVC-AQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-LRVRPPRLW 134
             |:| ||  |  |  :   | ||:||||||: |||||||||| |||   :  | |
Db      85 TALKGICRAQSEGRPCEYNSRIYQNGESFQPNCKHQCTCIDGAVGCIPLCPQELSLPNLG 144

Qy     135 CPHPRRVSIPGHCCEQWVCEDDA-KRPRKTAPRDTG---AFDAVGEVEAWHRN------- 183
           ||:|| | : | |||:|||: |: | | :       |    |||  |||    |
Db     145 CPNPRLVKVTGQCCEEWVCDQDSIKDPMEDQDGLLGKELGFDA-SEVELTRNNELIAVGK 203

Qy     184 ------------------------CIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQE 218
                                    ||  |: || || :|| |:|||::| | :|   :|
Db     204 GSSLKRLPVFGMEPRILYNPLQGQKCIVQTTSWSQCSKTCGTGISTRVTNDNPECRLVKE 263

Qy     219 SRLCNLRPCDVDIHTLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCC 278
           :|:| :|||    ::: :| ||||    :    : || |||:| : |:|||| |:| |||
Db     264 TRICEVRPCGQPVYSSLKKGKKCSKTKKSPEPVRFTYAGCLSVKKYRPKYCGSCVDGRCC 323

Qy     279 IPYKSKTIDVSFQCPDGLGFSRQVLWINACFCNLSCRNPND-------IFADLESYPD 329
```

```
          |  ::|: : |:| ||  ||: |: | :| || :| : |:       :| |:  : |
Db    324 TPQLTRTVKMRFRCEDGETFSKNVMMIQSCKCNYNCPHANEAAFPFYRLFNDIHKFRD 381


RESULT  15
Q9QZQ5
ID   Q9QZQ5      PRELIMINARY;      PRT;   351 AA.
AC   Q9QZQ5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   NOV PROTEIN.
GN   NOV.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY;
RX   MEDLINE=20035752; PubMed=10570975;
RA   Liu C., Liu X.-J., Crowe P.D., Kelner G.S., Fan J., Barry G., Manu F.,
RA   Ling N., De Souza E.B., Maki R.A.;
RT   "Nephroblastoma overexpressed gene (NOV) codes for a growth factor
RT   that induces protein tyrosine phosphorylation.";
RL   Gene 238:471-478(1999).
DR   EMBL; AF171936; AAD49371.1; -.
DR   InterPro; IPR000359; -.
DR   InterPro; IPR000867; -.
DR   InterPro; IPR000884; -.
DR   InterPro; IPR001007; -.
DR   Pfam; PF00007; Cys_knot; 1.
DR   Pfam; PF00090; tsp_1; 1.
DR   Pfam; PF00093; vwc; 1.
DR   Pfam; PF00219; IGFBP; 1.
DR   PROSITE; PS01185; CTCK_1; 1.
DR   PROSITE; PS01225; CTCK_2; 1.
DR   PROSITE; PS00222; IGF_BINDING; 1.
DR   PROSITE; PS01208; VWFC; UNKNOWN_1.
DR   SMART; SM00041; CT; 1.
SQ   SEQUENCE   351 AA;  38509 MW;  02619707DE7C1BFB CRC64;


  Query Match             37.1%;  Score 719.5;  DB 11;  Length 351;
  Best Local Similarity   41.6%;  Pred. No. 1.2e-63;
  Matches  134;  Conservative   46;  Mismatches  115;  Indels   27;  Gaps    7;

Qy      14 PQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAICDPHRGLYCDYSG 73
           |  |  | :||   | |  ||  : ||| || :||:| |::|:|   ||   ||||| |
Db      30 PSRC--PSQCPSISPTCAPGVRSVLDGCSCCPVCARQRGESCSEMRPCDQSSGLYCDRSA 87

Qy      74 DRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVGCTPLC-LRVRPPR 132
           | |    |:|    |  || ||| | ||: |:|||:|:||| || :|| | | | |  |
Db      88 D-PNNETGICMVPEGDNCVFDGVIYRNGEKFEPNCQYHCTCRDGQIGCVPRCQLDVLLPG 146

Qy     133 LWCPHPRRVSIPGHCCEQWVCEDDAK----------RPRKTAPRDTGAFDAVG-EVEAW 180
             || |::|::|| |||:| |  : |          ||  |          || |:
Db     147 PDCPAPKKVAVPGECCEKWTCGSEEKGTLGGLALPAYRPEAT----------VGVELSDS 196

Qy     181 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDI-HTLIKAGK 239
             |||  |: || || |||:|:|||::| | ||   :::||| :|||: :        ||
Db     197 SINCIEQTTEWSACSKSCGMGLSTRVTNRNLQCEMVKQTRLCMVRPCEQEPGEATDMKGK 256

Qy     240 KCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFS 299
           |||   :   |::     | |  :|:|::||:| | ||| |: :||| | |||  |
Db     257 KCLRTKKSLKSIHLQFKNCTSLYTYKPRFCGICSDGRCCTPFNTKTIQVEFQCLPGQIIK 316

Qy     300 RQVLWINACFCNLSCRNPNDIF 321
           : |: |  | |: :|   |: |
Db     317 KPVMVIGTCTCHSNCPQNNEAF 338
```

```
Search completed: July 26, 2001, 08:37:41
Job time: 188 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 26, 2001, 08:36:33 ; Search time 20.85 Seconds
(without alignments)
974.054 Million cell updates/sec

Title: US-09-325-019-2
Perfect score: 335
Sequence: 1 DFTPAPLEDTSSRPQFCKWP.....NPNDIFADLESYPDFSEIAN 335

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 7

Total number of hits satisfying chosen parameters: 96

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_0601:*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	335	100.0	335	21	AAY59247	Human connective t
2	335	100.0	345	20	AAY17640	Human putative mat
3	335	100.0	367	20	AAY17641	Human WISP-1 prote
4	335	100.0	367	22	AAB50975	Human PRO542 prote
5	234	69.9	345	20	AAY17642	Human WISP-1 varia
6	234	69.9	345	20	AAY17643	Human WISP-1 varia
7	234	69.9	367	20	AAY17644	Human WISP-1 varia
8	234	69.9	367	20	AAY17645	Human WISP-1 varia
9	165	49.3	345	20	AAY17652	Human WISP-1 varia
10	165	49.3	367	20	AAY17653	Human WISP-1 varia
11	52	15.5	345	20	AAY17646	Mouse putative mat
12	52	15.5	367	20	AAY17647	Mouse WISP-1 prote
13	10	3.0	17	18	AAW35962	Human monocyte mat
14	10	3.0	124	15	AAR46078	CYR61 like protein
15	10	3.0	379	13	AAR25565	Beta-IG-M1. Mus m
16	10	3.0	381	18	AAW35730	Human cysteine ric
17	10	3.0	381	18	AAW35957	Human monocyte mat
18	10	3.0	455	21	AAB43987	Human cancer assoc
19	8	2.4	122	21	AAG56333	Arabidopsis thalia
20	8	2.4	146	21	AAG56331	Arabidopsis thalia
21	8	2.4	227	20	AAY17648	Human putative mat
22	8	2.4	228	20	AAY17679	Human WISP-2 prote
23	8	2.4	229	20	AAY17678	Human WISP-2 prote
24	8	2.4	230	20	AAY17677	Human WISP-2 prote
25	8	2.4	231	20	AAY17676	Human WISP-2 prote
26	8	2.4	232	20	AAY17675	Human WISP-2 prote
27	8	2.4	233	20	AAY17674	Human WISP-2 prote
28	8	2.4	234	20	AAY17673	Human WISP-2 prote
29	8	2.4	235	20	AAY17672	Human WISP-2 prote
30	8	2.4	236	20	AAY17671	Human WISP-2 prote
31	8	2.4	237	20	AAY17670	Human WISP-2 prote
32	8	2.4	238	20	AAY17669	Human WISP-2 prote
33	8	2.4	239	20	AAY17668	Human WISP-2 prote
34	8	2.4	240	20	AAY17667	Human WISP-2 prote
35	8	2.4	241	20	AAY17666	Human WISP-2 prote
36	8	2.4	242	20	AAY17665	Human WISP-2 prote
37	8	2.4	243	20	AAY17664	Human WISP-2 prote
38	8	2.4	244	20	AAY17663	Human WISP-2 prote
39	8	2.4	245	20	AAY17662	Human WISP-2 prote
40	8	2.4	246	20	AAY17661	Human WISP-2 prote
41	8	2.4	247	20	AAY17660	Human WISP-2 prote
42	8	2.4	248	20	AAY17659	Human WISP-2 prote
43	8	2.4	249	20	AAY17658	Human WISP-2 prote
44	8	2.4	250	19	AAW37946	Human connective t
45	8	2.4	250	20	AAY17649	Human WISP-2 prote

ALIGNMENTS

RESULT 1
AAY59247
ID AAY59247 standard; Protein; 335 AA.
XX
AC AAY59247;
XX
DT 11-APR-2000 (first entry)
XX
DE Human connective tissue growth factor-4 (CTGF-4).
XX
KW Connective tissue growth factor-4; CTGF-4; human; immune system disorder;
KW hematopoietic disorder; autoimmune disorder; diabetes mellitus; asthma;
KW respiratory disorder; inflammation; hyperproliferative disorder;
KW infection; central nervous system disease; Alzheimer's disease; AIDS;
KW food additive.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 15..84
FT /note= "IGF binding domain"
FT Domain 28..36
FT /note= "conserved domain CD-I"
FT Domain 39..55
FT /note= "conserved domain CD-II"
FT Modified-site 54
FT /note= "potential N-glycosylation site"
FT Domain 61..70
FT /note= "conserved domain CD-III"
FT Domain 89..154
FT /note= "Von-Willebrand factor type C repeat fragment"
FT Domain 101..121
FT /note= "conserved domain CD-IV"
FT Modified-site 111

```
FT                  /note= "potential N-glycosylation site"
FT   Domain         144-154
FT                  /note= "conserved domain CD-V"
FT   Domain         184..228
FT                  /note= "sulfated glycoconjugate binding motif"
FT   Domain         194..213
FT                  /note= "conserved domain CD-VI"
FT   Domain         216..227
FT                  /note= "conserved domain CD-VII"
FT   Domain         236..241
FT                  /note= "conserved domain CD-VIII"
FT   Domain         241..316
FT                  /note= "C-terminal dimerisation and receptor-binding
FT                        domain"
FT   Modified-site  252
FT                  /note= "potential N-glycosylation site"
FT   Domain         253..260
FT                  /note= "conserved domain CD-IX"
FT   Domain         264..280
FT                  /note= "conserved domain CD-X"
FT   Domain         290..295
FT                  /note= "conserved domain CD-XI"
FT   Modified-site  311
FT                  /note= "potential N-glycosylation site"
XX
PN   WO9962927-A1.
XX
PD   09-DEC-1999.
XX
PF   03-JUN-1999;   99WO-US12150.
XX
PR   05-JUN-1998;   98US-0088320.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Young PE;
XX
DR   WPI; 2000-147042/13.
DR   N-PSDB; AAZ58613.
XX
PT   New isolated connective tissue growth factor-4, used for treating e.g.
PT   cancers -
XX
PS   Claim 11; Fig 1A-E; 196pp; English.
XX
CC   The invention provides an isolated human connective tissue growth factor
CC   -4 (CTGF-4) polypeptide. The CTGF-4 cDNA is deposited under ATCC No.
CC   209816. The CTGF-4 protein can be expressed by standard recombinant
CC   methodology. The polypeptides can be used for preventing, treating or
CC   ameliorating a medical condition . They may be useful in treating
CC   deficiencies or disorders of the immune system, by activating or
CC   inhibiting the proliferation, differentiation, or mobilization
CC   (chemotaxis) of immune cells, treating or detecting deficiencies or
CC   disorders of hematopoietic cells (e.g. blood protein disorders, ataxia
CC   telangiectasia, HIV infection, Di-george syndrome, anemia or
CC   hemoglobinuria), to modulate hemostatic (the stopping of bleeding) or
CC   thrombolytic activity (clot formation) (e.g. blood coagulation disorders,
CC   blood platelet disorders, or wounds resulting from trauma, or surgery),
CC   in treating or detecting autoimmune disorders (e.g. Addison's disease,
CC   rheumatoid arthritis, allergic encephalomyelitis, Goodpastures syndrome,
CC   multiple sclerosis, purpura, Reiter's disease, Guillain-Barre syndrome,
CC   systemic lupus erythematosus, insulin dependent diabetes mellitus or
CC   autoimmune inflammatory eye disease), treating asthma (particularly
CC   allergic asthma) or other respiratory problems (e.g. anaphylaxis,
CC   hypersensitivity to an antigenic molecule or blood group
CC   incompatibility), to treat and/or prevent organ rejection or graft-versus
CC   -host disease (GVHD), to modulate inflammation (septic shock, sepsis,
CC   arthritis, nephritis, cytokine or chemokine induced lung injury,
CC   inflammatory bowel disease, Crohn's disease, or resulting from over
CC   production of cytokines), to treat hyperproliferative disorders,
CC   including neoplasms in the abdomen, bone, breast, digestive system,
CC   liver, pancreas, peritoneum, endocrine glands, eye, head and neck,
CC   nervous (central and peripheral), lymphatic system, pelvic, skin, soft
CC   tissue, spleen, thoracic and urogenital, hypergammaglobulinemia,
CC   lymphoproliferative disorders, Waldenstron's macroglobulinemia,
CC   sarcoidosis), to treat or detect infectious agents, e.g. viruses (e.g.
CC   arthritis, bronchiollitis, encephalitis, eye infections, chronic fatigue
CC   syndrome, hepatitis, meningitis, AIDS, pneumonia, chickenpox, measles,
CC   mumps, parainfluenza, rabies, the common cold, polio, leukemia, rubella,
CC   sexually transmitted diseases, or skin diseases) bacterial or fungal
CC   agents (e.g. bacteremia, endocarditis, eye infections, gingivitis,
CC   opportunistic infections, respiratory tract infections, Lyme disease,
CC   cat-scratch disease, paratyphoid fever, food poisoning, pneumonia,
CC   gonorrhea and sexually transmitted diseases, meningitis, tuberculosis,
CC   lupus, gangrene, tetanus, rheumatic fever, urinary tract infections,
CC   wound infections), parasitic agents (e.g. scabies, dysentery, liver
CC   disease, malaria, toxoplasmosis), to differentiate, proliferate and
CC   attract cells, leading to the regeneration of tissues (e.g. repair,
CC   replace or protect tissue in wounds, burns, incisions or ulcers,
CC   osteoporosis, osteocarthritis, periodontal disease, liver failure,
CC   surgery, cosmetic plastic surgery, reperfusion injury) to proliferate and
CC   differentiate nerve cells ( e.g. spinal cord disorders, head trauma,
CC   cerebrovascular disease and stroke), localized neuropathies and central
CC   nervous system diseases (e.g. Alzheimer's disease, Parkinson's disease,
CC   Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager
CC   syndrome). They may also increase or decrease the differentiation or
CC   proliferation of embryonic stem cells and hematopoietic lineage, may be
CC   used to modulate mammalian characteristics such as body height, weight,
CC   hair color, eye color, skin, percentage of adipose tissue, pigmentation,
CC   size, and shape, to modulate mammalian metabolism affecting catabolism,
CC   anabolism, processing, utilization and storage of energy, to change a
CC   mammal's mental state or physical state by influencing biorhythms,
CC   caricadic rhythms, circadian rhythms, depression (including depressive
CC   disorders), tendency for violence, tolerance for pain, reproductive
CC   capabilities, hormonal or endocrine levels, appetite, libido, memory,
CC   stress, or other cognitive qualities, as a food additive or preservative,
CC   such as to increase or decrease storage capabilities, fat content, lipid,
CC   protein, carbohydrate, vitamins, minerals, cofactors or other nutritional
CC   components. Mutations in the PNs or the presence or amount of expression
CC   or activity of the polypeptides can be used for diagnosing a pathological
CC   condition or a susceptibility to a pathological condition. The CFTR-4
CC   polypeptides can also be used for identifying binding partners . The
CC   products can also be used for producing transgenic animals. The present
CC   sequence represents the CTGF-4 polypeptide.
XX
SQ   Sequence   335 AA;

  Query Match             100.0%;  Score 335;  DB 21;  Length 335;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  335;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 dftpapledtssrpqfckwpcecppspprcplgvslitdgceckmcaqqlgdnckeaai 60

Qy      61 CDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 cdphrglycdysgdrpryaigvcaqvvgvgcvldgvrynngqsfqpnckynctcidgavg 120

Qy     121 CTPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAW 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ctplclrvrpprlwcphprrvsipghcceqwvceddakrprktaprdtgafdavgeveaw 180

Qy     181 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 hrnciaytspwspcstscglgvstrisnvnaqcwpeqesrlcnlrpcdvdihtlikagkk 240

Qy     241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 clavyqpeasmnftlagcistrsyqpkycgvcmdnrccipyksktidvsfqcpdglgfsr 300

Qy     301 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
           |||||||||||||||||||||||||||||||||||
Db     301 qvlwinacfcnlscrnpndifadlesypdfseian 335
```

```
RESULT   2
AAY17640
ID   AAY17640 standard; Protein; 345 AA.
XX
AC   AAY17640;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Human putative mature WISP-1 protein SEQ ID NO:3.
XX
KW   WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW   connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW   leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW   tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW   kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW   connective tissue disorder; catabolic state; inflammation;
KW   testicular-related disorder; angiogenesis; immunological disorder.
XX
OS   Homo sapiens.
XX
PN   WO9921998-A1.
XX
PD   06-MAY-1999.
XX
PF   29-OCT-1998;   98WO-US22991.
XX
PR   14-APR-1998;   98US-0081695.
PR   29-OCT-1997;   97US-0063704.
PR   03-FEB-1998;   98US-0073612.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein DA,  Cohen RL,  Goddard A,  Gurney AL,  Hillan K;
PI   Lawrence DA,  Levine AJ,  Pennica D,  Roy MA,  Wood WI;
XX
DR   WPI; 1999-337420/28.
XX
PT   New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS   Claim 1; Page 162-163; 284pp; English.
XX
CC   The present invention describes Wnt-1 induced secreted polypeptides,
CC   WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC   and WISP-3 have homology to connective tissue growth factor (CTGF).
CC   Products from the present invention can be used to treat WISP-related
CC   disorders such as breast, ovarian, and colon cancer or melanoma. The
CC   products can be used to treat arteriosclerosis. The products can also be
CC   used to treat other diseases e.g. benign and malignant tumours,
CC   leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC   blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC   disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC   disorders, bone-related disorders such as osteoporosis, trauma such as
CC   burns, incisions, and other wounds, connective tissue disorders,
CC   catabolic states, testicular-related disorders, and inflammatory,
CC   angiogenic and immunologic disorders including arteriosclerosis. The
CC   products can also be used for detection and diagnosis especially of
CC   individuals with neoplastic cell growth or proliferation. The products
CC   can be used in the production of transgenic or knock-out animals.
CC   Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC   cells.
XX
SQ   Sequence   345 AA;

  Query Match           100.0%;  Score 335;  DB 20;  Length 345;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  335;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11 dftpapledtssrpqfckwpcecppspprcplgvslitdgcecckmcaqqlgdncteaai 70

Qy      61 CDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      71 cdphrglycdysgdrpryaigvcaqvvgvgcvldgvrynngqsfqpnckynctcidgavg 130

Qy     121 CTPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAW 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     131 ctplclrvrpprlwcphprrvsipghcceqwvceddakrprktaprdtgafdavgeveaw 190

Qy     181 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     191 hrnciaytspwspcstscglgvstrisnvnaqcwpeqesrlcnlrpcdvdihtlikagkk 250

Qy     241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251 clavyqpeasmnftlagcistrsyqpkycgvcmdnrccipyksktidvsfqcpdglgfsr 310

Qy     301 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
           ||||||||||||||||||||||||||||||||||
Db     311 qvlwinacfcnlscrnpndifadlesypdfseian 345


RESULT   3
AAY17641
ID   AAY17641 standard; Protein; 367 AA.
XX
AC   AAY17641;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Human WISP-1 protein SEQ ID NO:4.
XX
KW   WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW   connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW   leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW   tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW   kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW   connective tissue disorder; catabolic state; inflammation;
KW   testicular-related disorder; angiogenesis; immunological disorder.
XX
OS   Homo sapiens.
XX
PN   WO9921998-A1.
XX
PD   06-MAY-1999.
XX
PF   29-OCT-1998;   98WO-US22991.
XX
PR   14-APR-1998;   98US-0081695.
PR   29-OCT-1997;   97US-0063704.
PR   03-FEB-1998;   98US-0073612.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein DA,  Cohen RL,  Goddard A,  Gurney AL,  Hillan K;
PI   Lawrence DA,  Levine AJ,  Pennica D,  Roy MA,  Wood WI;
XX
DR   WPI; 1999-337420/28.
DR   N-PSDB; AAX76482.
XX
PT   New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS   Claim 4; Page 163-164; 284pp; English.
XX
CC   The present invention describes Wnt-1 induced secreted polypeptides,
CC   WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC   and WISP-3 have homology to connective tissue growth factor (CTGF).
CC   Products from the present invention can be used to treat WISP-related
CC   disorders such as breast, ovarian, and colon cancer or melanoma. The
CC   products can be used to treat arteriosclerosis. The products can also be
CC   used to treat other diseases e.g. benign and malignant tumours,
CC   leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal, and
```

```
CC   blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC   disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC   disorders, bone-related disorders such as osteoporosis, trauma such as
CC   burns, incisions, and other wounds, connective tissue disorders,
CC   catabolic states, testicular-related disorders, and inflammatory,
CC   angiogenic and immunologic disorders including arteriosclerosis. The
CC   products can also be used for detection and diagnosis especially of
CC   individuals with neoplastic cell growth or proliferation. The products
CC   can be used in the production of transgenic or knock-out animals.
CC   Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC   cells.
XX
SQ   Sequence   367 AA;


  Query Match           100.0%;  Score 335;  DB 20;  Length 367;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  335;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     33 dftpapledtssrpqfckwpcecppspprcplgvslitdgcecckmcaqqlgdncteaai 92

Qy     61 CDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     93 cdphrglycdysgdrpryaigvcaqvvgvgcvldgvrynngqsfqpnckynctcidgavg 152

Qy    121 CTPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAW 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    153 ctplclrvrpprlwcphprrvsipghcceqwvceddakrprktaprdtgafdavgeveaw 212

Qy    181 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    213 hrnciaytspwspcstscglgvstrisnvnaqcwpeqesrlcnlrpcdvdihtlikagkk 272

Qy    241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    273 clavyqpeasmnftlagcistrsyqpkycgvcmdnrccipyksktidvsfqcpdglgfsr 332

Qy    301 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
          |||||||||||||||||||||||||||||||||||
Db    333 qvlwinacfcnlscrnpndifadlesypdfseian 367


RESULT   4
AAB50975
ID   AAB50975 standard; Protein; 367 AA.
XX
AC   AAB50975;
XX
DT   21-MAR-2001  (first entry)
XX
DE   Human PRO542 protein.
XX
KW   Human; PRO; cytostatic; nootropic; neuroprotective; respiratory general;
KW   antiinflammatory; antiangiogenic; immunosuppressive; immunostimulant;
KW   PRO agonist; cancer; inflammatory disorder; immunological disorder.
XX
OS   Homo sapiens.
XX
PN   WO200073348-A2.
XX
PD   07-DEC-2000.
XX
PF   30-MAY-2000; 2000WO-US14941.
XX
PR   02-JUN-1999;   99WO-US12252.
PR   22-JUN-1999;   99US-0140650.
PR   23-JUN-1999;   99US-0141037.
PR   20-JUL-1999;   99US-0144758.
PR   01-SEP-1999;   99WO-US20111.
PR   08-SEP-1999;   99WO-US20594.
PR   29-OCT-1999;   99US-0162506.
PR   30-NOV-1999;   99WO-US28313.
PR   01-DEC-1999;   99WO-US28634.
PR   02-DEC-1999;   99WO-US28551.
PR   16-DEC-1999;   99WO-US30095.
PR   20-DEC-1999;   99WO-US30999.
PR   06-JAN-2000; 2000WO-US00376.
PR   11-FEB-2000; 2000WO-US03565.
PR   18-FEB-2000; 2000WO-US04341.
PR   18-FEB-2000; 2000WO-US04342.
PR   02-MAR-2000; 2000WO-US05841.
PR   03-MAR-2000; 2000US-0187202.
PR   10-MAR-2000; 2000WO-US06319.
PR   15-MAR-2000; 2000WO-US06884.
PR   30-MAR-2000; 2000WO-US08439.
PR   17-MAY-2000; 2000WO-US13705.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP,  Goddard A,  Gurney AL,  Hebert C,  Henzel W,  Kabakoff RC;
PI   Shelton DL,  Smith V,  Watanabe CK,  Wood WI;
XX
DR   WPI; 2001-016509/02.
DR   N-PSDB; AAC91577.
XX
PT   Twenty eight nucleic acids encoding PRO polypeptides which are useful
PT   for treating various tumors, e.g. breast cancer, and other
PT   inflammatory, angiogenic and immunological disorders -
XX
PS   Claim 31; Fig 50; 188pp; English.
XX
CC   The present sequence is one of twenty eight novel PRO polypeptides. The
CC   PRO polypeptides and their agonists, including antibodies, peptides, and
CC   small molecule agonists, may be used to treat various tumours, e.g.,
CC   cancers such as breast cancer, ovarian cancer, renal cancer, colorectal
CC   cancer, uterine cancer, prostate cancer, lung cancer, bladder cancer,
CC   central nervous system cancer, melanoma or leukaemia. They are also
CC   useful for treating other disorders such as neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal and
CC   blastocoelic disorders, and inflammatory, angiogenic and immunological
CC   disorders.
XX
SQ   Sequence   367 AA;


  Query Match           100.0%;  Score 335;  DB 22;  Length 367;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  335;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     33 dftpapledtssrpqfckwpcecppspprcplgvslitdgcecckmcaqqlgdncteaai 92

Qy     61 CDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     93 cdphrglycdysgdrpryaigvcaqvvgvgcvldgvrynngqsfqpnckynctcidgavg 152

Qy    121 CTPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAW 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    153 ctplclrvrpprlwcphprrvsipghcceqwvceddakrprktaprdtgafdavgeveaw 212

Qy    181 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    213 hrnciaytspwspcstscglgvstrisnvnaqcwpeqesrlcnlrpcdvdihtlikagkk 272

Qy    241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    273 clavyqpeasmnftlagcistrsyqpkycgvcmdnrccipyksktidvsfqcpdglgfsr 332

Qy    301 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
          |||||||||||||||||||||||||||||||||||
Db    333 qvlwinacfcnlscrnpndifadlesypdfseian 367
```

```
RESULT   5
AAY17642
ID   AAY17642 standard; Protein; 345 AA.
XX
AC   AAY17642;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Human WISP-1 variant protein SEQ ID NO:5.
XX
KW   WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW   connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW   leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW   tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW   kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW   connective tissue disorder; catabolic state; inflammation;
KW   testicular-related disorder; angiogenesis; immunological disorder.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9921998-A1.
XX
PD   06-MAY-1999.
XX
PF   29-OCT-1998;   98WO-US22991.
XX
PR   14-APR-1998;   98US-0081695.
PR   29-OCT-1997;   97US-0063704.
PR   03-FEB-1998;   98US-0073612.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein DA,  Cohen RL,  Goddard A,  Gurney AL,  Hillan K;
PI   Lawrence DA,  Levine AJ,  Pennica D,  Roy MA,  Wood WI;
XX
DR   WPI; 1999-337420/28.
XX
PT   New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS   Claim 5; Page 164-165; 284pp; English.
XX
CC   The present invention describes Wnt-1 induced secreted polypeptides,
CC   WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC   and WISP-3 have homology to connective tissue growth factor (CTGF).
CC   Products from the present invention can be used to treat WISP-related
CC   disorders such as breast, ovarian, and colon cancer or melanoma. The
CC   products can be used to treat arteriosclerosis. The products can also be
CC   used to treat other diseases e.g. benign and malignant tumours,
CC   leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC   blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC   disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC   disorders, bone-related disorders such as osteoporosis, trauma such as
CC   burns, incisions, and other wounds, connective tissue disorders,
CC   catabolic states, testicular-related disorders, and inflammatory,
CC   angiogenic and immunologic disorders including arteriosclerosis. The
CC   products can also be used for detection and diagnosis especially of
CC   individuals with neoplastic cell growth or proliferation. The products
CC   can be used in the production of transgenic or knock-out animals.
CC   Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC   cells.
XX
SQ   Sequence   345 AA;


  Query Match             69.9%;  Score 234;  DB 20;  Length 345;
  Best Local Similarity   99.7%;  Pred. No. 7.7e-228;
  Matches  334;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11 dftpapledtssrpqfckwpcecppspprcplgvslitdgcecckmcaqqlgdnctеaai 70

Qy      61 CDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      71 cdphrglycdysgdrpryaigvcaqvvgvgcvldgvrynngqsfqpnckynctcidgavg 130

Qy     121 CTPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAW 180
           |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db     131 ctplclrvrpprlwcphprrvsipghcceqwiceddakrprktaprdtgafdavgeveaw 190

Qy     181 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     191 hrnciaytspwspcstscglgvstrisnvnaqcwpeqesrlcnlrpcdvdihtlikagkk 250

Qy     241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251 clavyqpeasmnftlagcistrsyqpkycgvcmdnrccipyksktidvsfqcpdglgfsr 310

Qy     301 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
           |||||||||||||||||||||||||||||||||||
Db     311 qvlwinacfcnlscrnpndifadlesypdfseian 345


RESULT   6
AAY17643
ID   AAY17643 standard; Protein; 345 AA.
XX
AC   AAY17643;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Human WISP-1 variant protein SEQ ID NO:6.
XX
KW   WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW   connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW   leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW   tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW   kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW   connective tissue disorder; catabolic state; inflammation;
KW   testicular-related disorder; angiogenesis; immunological disorder.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9921998-A1.
XX
PD   06-MAY-1999.
XX
PF   29-OCT-1998;   98WO-US22991.
XX
PR   14-APR-1998;   98US-0081695.
PR   29-OCT-1997;   97US-0063704.
PR   03-FEB-1998;   98US-0073612.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein DA,  Cohen RL,  Goddard A,  Gurney AL,  Hillan K;
PI   Lawrence DA,  Levine AJ,  Pennica D,  Roy MA,  Wood WI;
XX
DR   WPI; 1999-337420/28.
XX
PT   New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS   Claim 5; Page 166-167; 284pp; English.
XX
CC   The present invention describes Wnt-1 induced secreted polypeptides,
CC   WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC   and WISP-3 have homology to connective tissue growth factor (CTGF).
CC   Products from the present invention can be used to treat WISP-related
CC   disorders such as breast, ovarian, and colon cancer or melanoma. The
CC   products can be used to treat arteriosclerosis. The products can also be
CC   used to treat other diseases e.g. benign and malignant tumours,
CC   leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal, and
```

```
CC   blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC   disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC   disorders, bone-related disorders such as osteoporosis, trauma such as
CC   burns, incisions, and other wounds, connective tissue disorders,
CC   catabolic states, testicular-related disorders, and inflammatory,
CC   angiogenic and immunologic disorders including arteriosclerosis. The
CC   products can also be used for detection and diagnosis especially of
CC   individuals with neoplastic cell growth or proliferation. The products
CC   can be used in the production of transgenic or knock-out animals.
CC   Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC   cells.
XX
SQ   Sequence   345 AA;


  Query Match             69.9%;  Score 234;  DB 20;  Length 345;
  Best Local Similarity   99.7%;  Pred. No. 7.7e-228;
  Matches  334;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11 dftpapledtssrpqfckwpcecppspprcplgvslitdgcecckmcaqqlgdncteaai 70

Qy      61 CDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      71 cdphrglycdysgdrpryaigvcaqvvgvgcvldgvrynngqsfqpnckynctcidgavg 130

Qy     121 CTPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAW 180
           |||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     131 ctplclrvrpprlwcphprrvsipghcceqwvceddakrprktaprdtgsfdavgeveaw 190

Qy     181 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     191 hrnciaytspwspcstscglgvstrisnvnaqcwpeqesrlcnlrpcdvdihtlikagkk 250

Qy     241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251 clavyqpeasmnftlagcistrsyqpkycgvcmdnrccipyksktidvsfqcpdglgfsr 310

Qy     301 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
           |||||||||||||||||||||||||||||||||||
Db     311 qvlwinacfcnlscrnpndifadlesypdfseian 345


RESULT   7
AAY17644
ID   AAY17644 standard; Protein; 367 AA.
XX
AC   AAY17644;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Human WISP-1 variant protein SEQ ID NO:7.
XX
KW   WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW   connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW   leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW   tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW   kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW   connective tissue disorder; catabolic state; inflammation;
KW   testicular-related disorder; angiogenesis; immunological disorder.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9921998-A1.
XX
PD   06-MAY-1999.
XX
PF   29-OCT-1998;   98WO-US22991.
XX
PR   14-APR-1998;   98US-0081695.
PR   29-OCT-1997;   97US-0063704.
PR   03-FEB-1998;   98US-0073612.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein DA,  Cohen RL,  Goddard A,  Gurney AL,  Hillan K;
PI   Lawrence DA,  Levine AJ,  Pennica D,  Roy MA,  Wood WI;
XX
DR   WPI; 1999-337420/28.
XX
PT   New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS   Claim 6; Page 167-168; 284pp; English.
XX
CC   The present invention describes Wnt-1 induced secreted polypeptides,
CC   WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC   and WISP-3 have homology to connective tissue growth factor (CTGF).
CC   Products from the present invention can be used to treat WISP-related
CC   disorders such as breast, ovarian, and colon cancer or melanoma. The
CC   products can be used to treat arteriosclerosis. The products can also be
CC   used to treat other diseases e.g. benign and malignant tumours,
CC   leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC   blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC   disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC   disorders, bone-related disorders such as osteoporosis, trauma such as
CC   burns, incisions, and other wounds, connective tissue disorders,
CC   catabolic states, testicular-related disorders, and inflammatory,
CC   angiogenic and immunologic disorders including arteriosclerosis. The
CC   products can also be used for detection and diagnosis especially of
CC   individuals with neoplastic cell growth or proliferation. The products
CC   can be used in the production of transgenic or knock-out animals.
CC   Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC   cells.
XX
SQ   Sequence   367 AA;


  Query Match             69.9%;  Score 234;  DB 20;  Length 367;
  Best Local Similarity   99.7%;  Pred. No. 8.1e-228;
  Matches  334;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      33 dftpapledtssrpqfckwpcecppspprcplgvslitdgcecckmcaqqlgdncteaai 92

Qy      61 CDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93 cdphrglycdysgdrpryaigvcaqvvgvgcvldgvrynngqsfqpnckynctcidgavg 152

Qy     121 CTPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAW 180
           |||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     153 ctplclrvrpprlwcphprrvsipghcceqwiceddakrprktaprdtgafdavgeveaw 212

Qy     181 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     213 hrnciaytspwspcstscglgvstrisnvnaqcwpeqesrlcnlrpcdvdihtlikagkk 272

Qy     241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     273 clavyqpeasmnftlagcistrsyqpkycgvcmdnrccipyksktidvsfqcpdglgfsr 332

Qy     301 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
           |||||||||||||||||||||||||||||||||||
Db     333 qvlwinacfcnlscrnpndifadlesypdfseian 367


RESULT   8
AAY17645
ID   AAY17645 standard; Protein; 367 AA.
XX
AC   AAY17645;
XX
DT   06-AUG-1999  (first entry)
```

```
XX
DE   Human WISP-1 variant protein SEQ ID NO:8.
XX
KW   WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW   connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW   leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW   tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW   kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW   connective tissue disorder; catabolic state; inflammation;
KW   testicular-related disorder; angiogenesis; immunological disorder.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9921998-A1.
XX
PD   06-MAY-1999.
XX
PF   29-OCT-1998;   98WO-US22991.
XX
PR   14-APR-1998;   98US-0081695.
PR   29-OCT-1997;   97US-0063704.
PR   03-FEB-1998;   98US-0073612.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein DA,  Cohen RL,  Goddard A,  Gurney AL,  Hillan K;
PI   Lawrence DA,  Levine AJ,  Pennica D,  Roy MA,  Wood WI;
XX
DR   WPI; 1999-337420/28.
XX
PT   New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS   Claim 6; Page 168-169; 284pp; English.
XX
CC   The present invention describes Wnt-1 induced secreted polypeptides,
CC   WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC   and WISP-3 have homology to connective tissue growth factor (CTGF).
CC   Products from the present invention can be used to treat WISP-related
CC   disorders such as breast, ovarian, and colon cancer or melanoma. The
CC   products can be used to treat arteriosclerosis. The products can also be
CC   used to treat other diseases e.g. benign and malignant tumours,
CC   leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC   blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC   disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC   disorders, bone-related disorders such as osteoporosis, trauma such as
CC   burns, incisions, and other wounds, connective tissue disorders,
CC   catabolic states, testicular-related disorders, and inflammatory,
CC   angiogenic and immunologic disorders including arteriosclerosis. The
CC   products can also be used for detection and diagnosis especially of
CC   individuals with neoplastic cell growth or proliferation. The products
CC   can be used in the production of transgenic or knock-out animals.
CC   Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC   cells.
XX
SQ   Sequence   367 AA;


  Query Match             69.9%;  Score 234;  DB 20;  Length 367;
  Best Local Similarity   99.7%;  Pred. No. 8.1e-228;
  Matches  334;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 DFTPAPLEDTSSRPQFCKWPCECPPSPPRCPLGVSLITDGCECCKMCAQQLGDNCTEAAI 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      33 dftpapledtssrpqfckwpcecppspprcplgvslitdgcecckmcaqqlgdncteaai 92

Qy      61 CDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRYNNGQSFQPNCKYNCTCIDGAVG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      93 cdphrglycdysgdrpryaigvcaqvvgvgcvldgvrynngqsfqpnckynctcidgavg 152

Qy     121 CTPLCLRVRPPRLWCPHPRRVSIPGHCCEQWVCEDDAKRPRKTAPRDTGAFDAVGEVEAW 180
           |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||

Db     153 ctplclrvrpprlwcphprrvsipghcceqwvceddakrprktaprdtgsfdavgeveaw 212

Qy     181 HRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVDIHTLIKAGKK 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     213 hrnciaytspwspcstscglgvstrisnvnaqcwpeqesrlcnlrpcdvdihtlikagkk 272

Qy     241 CLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSFQCPDGLGFSR 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     273 clavyqpeasmnftlagcistrsyqpkycgvcmdnrccipyksktidvsfqcpdglgfsr 332

Qy     301 QVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
           |||||||||||||||||||||||||||||||||||
Db     333 qvlwinacfcnlscrnpndifadlesypdfseian 367


RESULT   9
AAY17652
ID   AAY17652 standard; Protein; 345 AA.
XX
AC   AAY17652;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Human WISP-1 variant protein SEQ ID NO:21.
XX
KW   WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW   connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW   leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW   tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW   kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW   connective tissue disorder; catabolic state; inflammation;
KW   testicular-related disorder; angiogenesis; immunological disorder.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9921998-A1.
XX
PD   06-MAY-1999.
XX
PF   29-OCT-1998;   98WO-US22991.
XX
PR   14-APR-1998;   98US-0081695.
PR   29-OCT-1997;   97US-0063704.
PR   03-FEB-1998;   98US-0073612.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein DA,  Cohen RL,  Goddard A,  Gurney AL,  Hillan K;
PI   Lawrence DA,  Levine AJ,  Pennica D,  Roy MA,  Wood WI;
XX
DR   WPI; 1999-337420/28.
XX
PT   New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS   Claim 7; Page 182-183; 284pp; English.
XX
CC   The present invention describes Wnt-1 induced secreted polypeptides,
CC   WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC   and WISP-3 have homology to connective tissue growth factor (CTGF).
CC   Products from the present invention can be used to treat WISP-related
CC   disorders such as breast, ovarian, and colon cancer or melanoma. The
CC   products can be used to treat arteriosclerosis. The products can also be
CC   used to treat other diseases e.g. benign and malignant tumours,
CC   leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC   blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC   disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC   disorders, bone-related disorders such as osteoporosis, trauma such as
CC   burns, incisions, and other wounds, connective tissue disorders,
CC   catabolic states, testicular-related disorders, and inflammatory,
CC   angiogenic and immunologic disorders including arteriosclerosis. The
CC   products can also be used for detection and diagnosis especially of
```

```
CC   individuals with neoplastic cell growth or proliferation. The products
CC   can be used in the production of transgenic or knock-out animals.
CC   Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC   cells.
XX
SQ   Sequence   345 AA;


  Query Match            49.3%;  Score 165;  DB 20;  Length 345;
  Best Local Similarity  100.0%;  Pred. No. 3.1e-158;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     171 FDAVGEVEAWHRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVD 230
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 fdavgeveawhrnciaytspwspcstscglgvstrisnvnaqcwpeqesrlcnlrpcdvd 240

Qy     231 IHTLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSF 290
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ihtlikagkkclavyqpeasmnftlagcistrsyqpkycgvcmdnrccipyksktidvsf 300

Qy     291 QCPDGLGFSRQVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
           |||||||||||||||||||||||||||||||||||||||||||||
Db     301 qcpdglgfsrqvlwinacfcnlscrnpndifadlesypdfseian 345


RESULT  10
AAY17653
ID   AAY17653 standard; Protein; 367 AA.
XX
AC   AAY17653;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Human WISP-1 variant protein SEQ ID NO:22.
XX
KW   WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW   connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW   leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW   tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW   kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW   connective tissue disorder; catabolic state; inflammation;
KW   testicular-related disorder; angiogenesis; immunological disorder.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9921998-A1.
XX
PD   06-MAY-1999.
XX
PF   29-OCT-1998;   98WO-US22991.
XX
PR   14-APR-1998;   98US-0081695.
PR   29-OCT-1997;   97US-0063704.
PR   03-FEB-1998;   98US-0073612.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein DA,  Cohen RL,  Goddard A,  Gurney AL,  Hillan K;
PI   Lawrence DA,  Levine AJ,  Pennica D,  Roy MA,  Wood WI;
XX
DR   WPI; 1999-337420/28.
XX
PT   New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS   Claim 7; Page 183-184; 284pp; English.
XX
CC   The present invention describes Wnt-1 induced secreted polypeptides,
CC   WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC   and WISP-3 have homology to connective tissue growth factor (CTGF).
CC   Products from the present invention can be used to treat WISP-related
CC   disorders such as breast, ovarian, and colon cancer or melanoma. The
CC   products can be used to treat arteriosclerosis. The products can also be
CC   used to treat other diseases e.g. benign and malignant tumours,
CC   leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC   blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC   disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC   disorders, bone-related disorders such as osteoporosis, trauma such as
CC   burns, incisions, and other wounds, connective tissue disorders,
CC   catabolic states, testicular-related disorders, and inflammatory,
CC   angiogenic and immunologic disorders including arteriosclerosis. The
CC   products can also be used for detection and diagnosis especially of
CC   individuals with neoplastic cell growth or proliferation. The products
CC   can be used in the production of transgenic or knock-out animals.
CC   Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC   cells.
XX
SQ   Sequence   367 AA;


  Query Match            49.3%;  Score 165;  DB 20;  Length 367;
  Best Local Similarity  100.0%;  Pred. No. 3.3e-158;
  Matches  165;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     171 FDAVGEVEAWHRNCIAYTSPWSPCSTSCGLGVSTRISNVNAQCWPEQESRLCNLRPCDVD 230
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 fdavgeveawhrnciaytspwspcstscglgvstrisnvnaqcwpeqesrlcnlrpcdvd 262

Qy     231 IHTLIKAGKKCLAVYQPEASMNFTLAGCISTRSYQPKYCGVCMDNRCCIPYKSKTIDVSF 290
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     263 ihtlikagkkclavyqpeasmnftlagcistrsyqpkycgvcmdnrccipyksktidvsf 322

Qy     291 QCPDGLGFSRQVLWINACFCNLSCRNPNDIFADLESYPDFSEIAN 335
           |||||||||||||||||||||||||||||||||||||||||||||
Db     323 qcpdglgfsrqvlwinacfcnlscrnpndifadlesypdfseian 367


RESULT  11
AAY17646
ID   AAY17646 standard; Protein; 345 AA.
XX
AC   AAY17646;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Mouse putative mature WISP-1 protein SEQ ID NO:11.
XX
KW   WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW   connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW   leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW   tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW   kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW   connective tissue disorder; catabolic state; inflammation;
KW   testicular-related disorder; angiogenesis; immunological disorder.
XX
OS   Mus sp.
XX
PN   WO9921998-A1.
XX
PD   06-MAY-1999.
XX
PF   29-OCT-1998;   98WO-US22991.
XX
PR   14-APR-1998;   98US-0081695.
PR   29-OCT-1997;   97US-0063704.
PR   03-FEB-1998;   98US-0073612.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein DA,  Cohen RL,  Goddard A,  Gurney AL,  Hillan K;
PI   Lawrence DA,  Levine AJ,  Pennica D,  Roy MA,  Wood WI;
XX
DR   WPI; 1999-337420/28.
XX
PT   New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
```

```
XX
PS   Claim 9; Page 172-173; 284pp; English.
XX
CC   The present invention describes Wnt-1 induced secreted polypeptides,
CC   WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC   and WISP-3 have homology to connective tissue growth factor (CTGF).
CC   Products from the present invention can be used to treat WISP-related
CC   disorders such as breast, ovarian, and colon cancer or melanoma. The
CC   products can be used to treat arteriosclerosis. The products can also be
CC   used to treat other diseases e.g. benign and malignant tumours,
CC   leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC   blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC   disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC   disorders, bone-related disorders such as osteoporosis, trauma such as
CC   burns, incisions, and other wounds, connective tissue disorders,
CC   catabolic states, testicular-related disorders, and inflammatory,
CC   angiogenic and immunologic disorders including arteriosclerosis. The
CC   products can also be used for detection and diagnosis especially of
CC   individuals with neoplastic cell growth or proliferation. The products
CC   can be used in the production of transgenic or knock-out animals.
CC   Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC   cells.
XX
SQ   Sequence   345 AA;


  Query Match             15.5%;  Score 52;  DB 20;  Length 345;
  Best Local Similarity   100.0%;  Pred. No. 3e-44;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      47 CAQQLGDNCTEAAICDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRY 98
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57 caqqlgdncteaaicdphrglycdysgdrpryaigvcaqvvgvgcvldgvry 108


RESULT  12
AAY17647
ID   AAY17647 standard; Protein; 367 AA.
XX
AC   AAY17647;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Mouse WISP-1 protein SEQ ID NO:12.
XX
KW   WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
KW   connective tissue growth factor; cancer; melanoma; arteriosclerosis;
KW   leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
KW   tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
KW   kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
KW   connective tissue disorder; catabolic state; inflammation;
KW   testicular-related disorder; angiogenesis; immunological disorder.
XX
OS   Mus sp.
XX
PN   WO9921998-A1.
XX
PD   06-MAY-1999.
XX
PF   29-OCT-1998;   98WO-US22991.
XX
PR   14-APR-1998;   98US-0081695.
PR   29-OCT-1997;   97US-0063704.
PR   03-FEB-1998;   98US-0073612.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein DA,  Cohen RL,  Goddard A,  Gurney AL,  Hillan K;
PI   Lawrence DA,  Levine AJ,  Pennica D,  Roy MA,  Wood WI;
XX
DR   WPI; 1999-337420/28.
DR   N-PSDB; AAX76484.
XX
PT   New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
PS   Claim 9; Page 173-174; 284pp; English.
XX
CC   The present invention describes Wnt-1 induced secreted polypeptides,
CC   WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC   and WISP-3 have homology to connective tissue growth factor (CTGF).
CC   Products from the present invention can be used to treat WISP-related
CC   disorders such as breast, ovarian, and colon cancer or melanoma. The
CC   products can be used to treat arteriosclerosis. The products can also be
CC   used to treat other diseases e.g. benign and malignant tumours,
CC   leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC   blastocoelic disorders, haematopoiesis-related disorders, tissue-growth
CC   disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC   disorders, bone-related disorders such as osteoporosis, trauma such as
CC   burns, incisions, and other wounds, connective tissue disorders,
CC   catabolic states, testicular-related disorders, and inflammatory,
CC   angiogenic and immunologic disorders including arteriosclerosis. The
CC   products can also be used for detection and diagnosis especially of
CC   individuals with neoplastic cell growth or proliferation. The products
CC   can be used in the production of transgenic or knock-out animals.
CC   Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC   cells.
XX
SQ   Sequence   367 AA;


  Query Match             15.5%;  Score 52;  DB 20;  Length 367;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-44;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      47 CAQQLGDNCTEAAICDPHRGLYCDYSGDRPRYAIGVCAQVVGVGCVLDGVRY 98
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      79 caqqlgdncteaaicdphrglycdysgdrpryaigvcaqvvgvgcvldgvry 130


RESULT  13
AAW35962
ID   AAW35962 standard; peptide; 17 AA.
XX
AC   AAW35962;
XX
DT   05-MAR-1998  (first entry)
XX
DE   Human monocyte mature differentiation factor peptide fragment.
XX
KW   Human; monocyte; mature; differentiation factor; MMDF; macrophage;
KW   cancer; immune activator; tissue culture; infectious disease.
XX
OS   Homo sapiens.
XX
PN   JP09234079-A.
XX
PD   09-SEP-1997.
XX
PF   04-MAR-1996;   96JP-0075236.
XX
PR   04-MAR-1996;   96JP-0075236.
XX
PA   (TOYM ) TOYOBO KK.
XX
DR   WPI; 1997-497320/46.
XX
PT   A monocyte mature differentiation factor - useful for the long term
PT   tissue culture of macrophage(s)
XX
PS   Example 5; Page 15; 22pp; Japanese.
XX
CC   The present sequence represents a peptide fragment from a monocyte
CC   mature differentiation factor (MMDF) (full length shown on AAW35957)
CC   which maintains the life of macrophages for long periods in liquid
```

```
CC   culture. MMDF can be used as an anti-cancer agent, an immune
CC   activator and to treat infectious diseases.
XX
SQ   Sequence   17 AA;


  Query Match             3.0%;  Score 10;  DB 18;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 0.0056;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    112 CTCIDGAVGC 121
          ||||||||||
Db      2 ctcidgavgc 11


RESULT  14
AAR46078
ID   AAR46078 standard; Protein; 124 AA.
XX
AC   AAR46078;
XX
DT   19-OCT-1994  (first entry)
XX
DE   CYR61 like protein.
XX
KW   Human cDNA; library; enzyme; protein.
XX
OS   Homo sapiens.
XX
PN   WO9403599-A.
XX
PD   17-FEB-1994.
XX
PF   04-AUG-1993;   93WO-JP01095.
XX
PR   04-AUG-1992;   92JP-0208077.
PR   13-NOV-1992;   92JP-0327619.
PR   26-FEB-1993;   93JP-0061431.
XX
PA   (SAGA ) SAGAMI CHEM RES CENTRE.
XX
PI   Iwahori A,  Kato S,  Kato T,  Kim N,  Oh S,  Sekine S;
XX
DR   WPI; 1994-065688/08.
DR   N-PSDB; AAQ57417.
XX
PT   cDNA of human origin and proteins coded by it - which may be
PT   expressed by in vivo or in vitro translation using sense RNA or
PT   antisense DNA corresponding to the cDNA.
XX
PS   Claim 1; Page 32-33; 167pp; Japanese.
XX
CC   mRNA expressed in human fibrosarcoma cell line HT-1080 was
CC   isolated and used to construct a cDNA library using vector
CC   pKA1. Clone HP00021 encoding CYR61-like protein
CC   was isolated.
XX
SQ   Sequence   124 AA;


  Query Match             3.0%;  Score 10;  DB 15;  Length 124;
  Best Local Similarity   100.0%;  Pred. No. 0.03;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    112 CTCIDGAVGC 121
          ||||||||||
Db     38 ctcidgavgc 47


RESULT  15
AAR25565
ID   AAR25565 standard; Protein; 379 AA.
XX
AC   AAR25565;
XX
DT   18-JAN-1993  (first entry)
XX
DE   Beta-IG-M1.
XX
KW   Transforming growth factor beta; induced; CEF-10; v-src; chicken;
KW   embryo; fibroblasts; TGF-beta.
XX
OS   Mus musculus.
XX
PN   EP495674-A.
XX
PD   22-JUL-1992.
XX
PF   17-JAN-1992;   92EP-0300429.
XX
PR   18-JAN-1991;   91US-0642991.
PR   10-JAN-1992;   92US-0816270.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Brunner AM,  Chinn J,  Neubauer MG,  Purchio AF;
XX
DR   WPI; 1992-243508/30.
DR   N-PSDB; AAQ26421.
XX
PT   TGF-beta induced gene family - encodes proteins involved in
PT   growth and differentiation effects of TGF-beta-1
XX
PS   Claim 2; Fig 1; 35pp; English.
XX
CC   The protein sequence was deduced from the DNA sequence obtd. by
CC   screening a cDNA library made from AKR-2B mouse cells induced with
CC   TGF-beta1 and cyclohexamide with two probes from untreated AKR-2B
CC   mRNA and AKR-2B mRNA from cells treated with cyclohexamide and TGF-
CC   beta1.  The proteins encoded by hybridising colonies (beta-IG-M1 and
CC   beta-IG-M2) contain 38 Cys residues and are induced by TGF-beta1.
CC   Beta-IG-M1 displays 80 percent homology to the CEF-10 protein
CC   induced by v-src in chicken embryo fibroblasts and is identical
CC   to the protein encoded by cyr61, an immediate early response gene
CC   induced in quiescent BALB 3T3 cells by serum treatment.  Residues
CC   49-56 of beta-IG-M1 conform to the GCGCCXXC motif reported in the
CC   amino half of insulin-like growth factor (IGF) binding proteins.
CC   The C-terminal Cys rich region of beta-IG-M1, -M2 and CEF-10 contain
CC   an amino acid sequence with strong homology to a motif found near the
CC   C-terminal of the malarial circumsporozoite (CS) protein, which is
CC   highly conserved among all species of malarial parasites sequenced
CC   to date (designated region II).  This motif is also found in
CC   other proteins which have cell adhesive properties that mediate
CC   cell-cell and cell-extracellular matrix interactions, such as
CC   properdin, thrombospondin, and TRAP.  The proteins encoded by
CC   TGF-beta induced genes are likely to be involved in mediation of
CC   the biological effects of TGF-beta relating to cell growth and
CC   differentiation.   See also AAR25566.
XX
SQ   Sequence   379 AA;


  Query Match             3.0%;  Score 10;  DB 13;  Length 379;
  Best Local Similarity   100.0%;  Pred. No. 0.076;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    112 CTCIDGAVGC 121
          ||||||||||
Db    121 ctcidgavgc 130


Search completed: July 26, 2001, 08:38:27
Job time: 114 sec
```